UNIVERSITEIT
GENT

Faculteit Wetenschappen
Vakgroep Zuivere Wiskunde en Computeralgebra

The Reverse Mathematics of Elementary Recursive Nonstandard Analysis: A Robust Contribution to the Foundations of Mathematics

Sam Sanders

Promotoren: Christian Impens
Andreas Weiermann

Proefschrift voorgelegd aan de Faculteit Wetenschappen tot het behalen van de graad van Doctor in de Wetenschappen richting Wiskunde.

I dedicate this dissertation to the founding fathers of Reverse Mathematics and Nonstandard Analysis.

Contents

Preface

Reverse Mathematics (RM) is a program in the Foundations of Mathematics founded by Harvey Friedman in the Seventies ([**17**, **18**]). The aim of RM is to determine the minimal axioms required to prove a certain theorem of 'ordinary' mathematics. In many cases one observes that these minimal axioms are also equivalent to this theorem. This phenomenon is called the 'Main Theme' of RM and theorem 1.2 is a good example thereof. In practice, most theorems of everyday mathematics are equivalent to one of the four systems WKL_0, ACA_0, ATR_0 and Π^1_1-CA_0 or provable in the base theory RCA_0. An excellent introduction to RM is Stephen Simpson's monograph [**46**]. Nonstandard Analysis has always played an important role in RM. ([**32**, **52**, **53**]).

One of the open problems in the literature is the RM of theories of first-order strength $I\Delta_0 + \exp$ ([**46**, p. 406]). In Chapter I, we formulate a solution to this problem in theorem 1.3. This theorem shows that many of the equivalences from theorem 1.2 remain correct when we replace equality by infinitesimal proximity '$\approx$' from Nonstandard Analysis. The base theory now is ERNA, a nonstandard extension of $I\Delta_0 + \exp$. The principle that corresponds to 'Weak König's lemma' is the Universal Transfer Principle (see axiom schema 1.57). In particular, one can say that the RM of ERNA$+\Pi_1$-TRANS is a 'copy up to infinitesimals' of the RM of WKL_0. This implies that RM is 'robust' in the sense this term is used in Statistics and Computer Science ([**25**, **35**]).

Furthermore, we obtain applications of our results in Theoretical Physics in the form of the 'Isomorphism Theorem' (see theorem 1.106). This philosophical excursion is the first application of RM outside of Mathematics and implies that 'whether reality is continuous or discrete is undecidable because of the way mathematics is used in Physics' (see paragraph 3.2.4, p. 53). We briefly explore a connection with the program 'Constructive Reverse Mathematics' ([**30**, **31**]) and in the rest of Chapter I, we consider the RM of ACA_0 and related systems. In particular, we prove theorem 1.161, which is a first step towards a 'copy up to infinitesimals' of the RM of ACA_0. However, one major aesthetic problem with these results is the introduction of extra quantifiers in many of the theorems listed in theorem 1.3 (see e.g. theorem 1.94). To overcome this hurdle, we explore Relative Nonstandard Analysis in Chapters II and III. This new framework involves many degrees of infinity instead of the classical 'binary' picture where only two degrees 'finite' and 'infinite' are available. We extend ERNA to a theory of Relative Nonstandard Analysis called $\mathrm{ERNA}^{\mathbb{A}}$ and show how this theory and its extensions allow for a completely quantifier-free development of analysis. We also study the metamathematics of $\mathrm{ERNA}^{\mathbb{A}}$, motivated by RM. Several ERNA-theorems would not have been discovered without considering $\mathrm{ERNA}^{\mathbb{A}}$ first.

CHAPTER I

ERNA and Reverse Mathematics

That through which all things come
into being, is not a thing in itself.

Tao Te Ching
LAO TSE

1. Introduction

1.1. Introducing ERNA. *Hilbert's Program*, proposed in 1921, called for an axiomatic formalization of mathematics, together with a proof that this axiomatization is consistent. The consistency proof itself was to be carried out using only what Hilbert called *finitary* methods. In due time, many characterized Hilbert's informal notion of 'finitary' as that which can be formalized in Primitive Recursive Arithmetic (PRA), proposed in 1923 by Skolem (see e.g. [**51**]).

By Gödel's second incompleteness theorem (1931) it became evident that only *partial* realizations of Hilbert's program are possible. The system proposed by Chuaqui and Suppes, recently adapted by Rössler and Jeřábek, is such a partial realization, in that it provides an axiomatic foundation for basic analysis, with a PRA consistency proof ([**11**], [**42**]). Sommer and Suppes's improved system allows definition by recursion, which does away with a lot of explicit axioms, and still has a PRA proof of consistency ([**49**, p. 21]). This system is called *Elementary Recursive Nonstandard Analysis*, in short ERNA. Its consistency is proved via Herbrand's Theorem (1930), which is restricted to quantifier-free formulas $Q(x_1,\dots,x_n)$, usually containing free variables. Alternatively, one might say it is restricted to universal sentences

$$(\forall x_1)\dots(\forall x_n)Q(x_1,\dots,x_n).$$

We will use Herbrand's theorem in the following form (see [**11**] and [**49**]); for more details, see [**8**] and [**21**].

1.1. THEOREM. *A quantifier-free theory T is consistent if and only if every finite set of instantiated axioms of T is consistent.*

Since Herbrand's theorem requires that ERNA's axioms be written in a quantifier-free form, some axioms definitely look artificial. Fortunately, theorems do not suffer from the quantifier-free restriction.

As it turns out, ERNA is not strong enough to develop basic analysis (see theorem 1.3) and hence an extension of ERNA is required. In section 2.2, we extend ERNA with a Transfer principle for universal sentences, called Π_1-TRANS. In nonstandard mathematics, Transfer expresses Leibniz's principle that the 'same' laws hold for standard and nonstandard objects alike. The consistency of the extended theory

ERNA + Π_1-TRANS$^-$ is provable in PRA via a finite iteration of ERNA's consistency proof (see theorem 1.58). The theory ERNA + Π_1-TRANS has important applications in 'Reverse Mathematics', introduced next.

1.2. Introducing Reverse Mathematics. Reverse Mathematics is a program in Foundations of Mathematics founded around 1975 by Harvey Friedman ([**17**] and [**18**]) and developed intensely by Stephen Simpson, Kazuyuki Tanaka and others; for an overview of the subject, see [**46**] and [**47**]. The goal of Reverse Mathematics is to determine what (minimal) axiom system is *necessary* to prove a particular theorem of ordinary mathematics. By now, it is well known that large portions of mathematics (especially so in analysis) can be carried out in systems far weaker than ZFC, the 'usual' background theory for mathematics. Classifying theorems according to their logical strength reveals the following striking phenomenon: '*It turns out that, in many particular cases, if a mathematical theorem is proved from appropriately weak set existence axioms, then the axioms will be logically equivalent to the theorem*' ([**46**, Preface]). This recurring phenomenon is called the 'Main theme' of Reverse Mathematics (see e.g. [**45**]) and a good instance is the following theorem from [**46**, p. 36].

1.2. Theorem (Reverse Mathematics for WKL_0). *Within* RCA_0*, one can prove that Weak König's Lemma* (WKL) *is equivalent to each of the following mathematical statements:*

(1) *The Heine-Borel lemma: every covering of* $[0,1]$ *by a sequence of open intervals has a finite subcovering.*
(2) *Every covering of a compact metric space by a sequence of open sets has a finite subcovering.*
(3) *Every continuous real-valued function on* $[0,1]$*, or on any compact metric space, is bounded.*
(4) *Every continuous real-valued function on* $[0,1]$*, or on any compact metric space, is uniformly continuous.*
(5) *Every continuous real-valued function on* $[0,1]$ *is Riemann integrable.*
(6) *The maximum principle: every continuous real-valued function on* $[0,1]$*, or on any compact metric space, is bounded, or* (*equivalently*) *has a supremum or* (*equivalently*) *attains its maximum.*
(7) *The Peano existence theorem: if* $f(x,y)$ *is continuous in the neighbourhood of* $(0,0)$*, then the initial value problem* $y' = f(x,y)$*,* $y(0) = 0$ *has a continuously differentiable solution in the neighbourhood of* $(0,0)$*.*
(8) *Gödel's completeness theorem: every at most countable consistent set of sentences in the predicate calculus has a countable model.*
(9) *Every countable commutative ring has a prime ideal.*
(10) *Every countable field* (*of characteristic* 0) *has a unique algebraic closure.*
(11) *Every countable formally real field is orderable.*
(12) *Every countable formally real field has a* (*unique*) *real closure.*
(13) *Brouwer's fixed point theorem: every uniformly continuous function from* $[0,1]^n$ *to* $[0,1]^n$ *has a fixed point.*
(14) *The separable Hahn-Banach theorem: if* f *is a bounded linear functional on a subspace of a separable Banach space, and if* $\|f\| \leq 1$*, then* f *has an extension* $\tilde{f}$ *to the whole space such that* $\|\tilde{f}\| \leq 1$*.*

Below, we will establish a similar theorem for ERNA. For future reference, we list some of the arguments pointing in this direction. First, by theorem 1.74, there is an interpretation of WKL_0 in $\mathrm{ERNA} + \Pi_1$-TRANS. Hence, it is to be expected that some of the equivalent formulations of WKL have an interpretation in ERNA too. Second, in [**32**], Keisler introduces a nonstandard and conservative extension of WKL_0, called ${}^*\mathrm{WKL}_0$. It is defined as ${}^*\Sigma\mathrm{PA} + \mathrm{STP}$, where ${}^*\Sigma\mathrm{PA}$ is a nonstandard theory and STP is the second-order principle that any set of naturals can be coded into a hyperinteger and vice versa. As part of STP plays the role of WKL, other nonstandard principles, like Π_1-TRANS may have similar properties.

Third, ERNA can prove results of basic analysis 'up to infinitesimals'; see e.g. [**50**], where the proof of ERNA's version of the above item (7) is outlined. This suggests that replacing equality with equality up to infinitesimals might translate some of the equivalences in theorem 1.2 into ERNA. Fourth, Σ_1-separation for subsets of $\mathbb{N}$ is provable in $\mathrm{ERNA} + \Pi_1$-TRANS (see theorem 1.104). The former schema is equivalent to WKL ([**46**, IV.4.4]). Fifth, in [**46**, Remark X.4.3] Simpson suggests reconsidering the results of Reverse Mathematics for WKL_0 in the weaker theory WKL_0^*. For ERNA, which has roughly the same first-order strength, we will prove the following theorem; it contains several statements, translated from theorem 1.2 and [**46**] into ERNA's language, while preserving equivalence. For the definitions, see below.

1.3. Theorem (Reverse Mathematics for $\mathrm{ERNA}+\Pi_1$-TRANS). *The theory* ERNA *proves the equivalence between* Π_1-TRANS *and each of the following theorems concerning near-standard functions:*

(1) *Every S-continuous function on* $[0,1]$*, or on any interval, is bounded.*
(2) *Every S-continuous function on* $[0,1]$*, or on any interval, is continuous there.*
(3) *Every S-continuous function on* $[0,1]$*, or on any interval, is Riemann integrable.*
(4) *Weierstraß' theorem: every S-continuous function on* $[0,1]$*, or on any interval, has, or attains a supremum, up to infinitesimals.*
(5) *The uniform Brouwer fixed point theorem: every S-continuous function* $\phi : [0,1] \to [0,1]$ *has a fixed point up to infinitesimals of arbitrary depth.*
(6) *The first fundamental theorem of calculus.*
(7) *The Peano existence theorem for ordinary differential equations.*
(8) *The Cauchy completeness, up to infinitesimals, of* ERNA*'s field.*
(9) *Every S-continuous function on* $[0,1]$ *has a modulus of uniform continuity.*
(10) *The Weierstraß approximation theorem.*

A common feature of the items in the theorem is that strict equality has been replaced with $\approx$, i.e. equality up to infinitesimals. This seems the price to be paid for 'pushing down' into ERNA the theorems equivalent to WKL. For instance, item (5) guarantees that there is a number x_0 in $[0,1]$ such that $\phi(x_0) \approx x_0$, i.e. a fixpoint up to infinitesimals, but in general there is no point x_1 such that $\phi(x_1) = x_1$. In this way, one could say that the Reverse Mathematics of $\mathrm{ERNA} + \Pi_1$-TRANS is a 'copy up to infinitesimals' of the Reverse Mathematics of WKL_0.

Below, we prove theorem 1.3 in ERNA and briefly explore a possible connection between the Reverse Mathematics of ERNA$+\Pi_1$-TRANS and the program of Constructive Reverse Mathematics. We also demonstrate that our results have implications for Physics in the form of the Isomorphism Theorem (see paragraph 3.2.4). In particular, we show that 'whether reality is continuous or discrete is undecidable because of the way mathematics is used in Physics'. Recently, the question has arisen whether Reverse Mathematics has implications outside mathematics and, to the best of our knowledge, we have obtained the first example. To conclude this chapter, we study the Reverse Mathematics of ACA_0 in ERNA. Although work is still in progress, we obtain theorem 1.161, which is the analog of theorem 1.3 for part of the Reverse Mathematics of ACA_0. We also answer several open questions of Avigad from [**1**] with regard to Reverse Mathematics and Nonstandard Analysis.

Furthermore, theorems 1.3 and 1.161 imply that Reverse Mathematics is 'robust' in the sense of computer science and statistics. Indeed, in mathematics, the branch 'robust statistics' attempts to 'produce estimators that are not particularly affected by small departures from model assumptions' ([**25**]), i.e. the methods should be reasonably resistant to errors in the results, produced by deviations from assumptions. In computer science, the word robust 'refers to an operating system or other program that performs well not only under ordinary conditions but also under unusual conditions that stress its designers' assumptions' ([**35**]). In this way, by theorem 1.3, the Reverse Mathematics of WKL_0 is 'robust with respect to infinitesimal error'. Alternatively, the Reverse Mathematics of WKL_0 can be seen as an idealisation of that of ERNA$+\Pi_1$-TRANS, where the latter corresponds better to physical reality.

2. ERNA, the system

In this section we describe ERNA and its fundamental features. Undocumented results are quoted from [**49**].

1.4. Notation. *$\mathbb{N} = \{0, 1, 2, \dots\}$ consists of the (finite) nonnegative integers.*

1.5. Notation. *$\vec{x}$ stands for some finite (possibly empty) sequence $(x_1, \dots, x_k)$.*

1.6. Notation. *$\tau(\vec{x})$ denotes a term in which $\vec{x} = (x_1, \dots, x_k)$ is the list of the distinct free variables.*

2.1. Language and axioms.

2.1.1. *The language of ERNA.*

- connectives: $\wedge, \neg, \vee, \rightarrow, \leftrightarrow$
- quantifiers: $\forall, \exists$
- an infinite set of variables
- relation symbols:[1]
  - binary $x = y$
  - binary $x \leq y$
  - unary $\mathcal{I}(x)$, read as 'x is infinitesimal', also written '$x \approx 0$'
  - unary $\mathcal{N}(x)$, read as 'x is hypernatural'.
- individual constant symbols:
  - 0
  - 1

[1]For better readibility we express unary relations in x and binary ones in (x, y).

– ε (Axiom 1.14.(6) asserts that ε is a positive infinitesimal hyperrational.)
– ω (The axioms 1.14.(7) and 1.10.(4) assert that $\omega = 1/\varepsilon$ is an infinite hypernatural.)
– $\uparrow$, to be read as 'undefined'.

1.7. Notation. '*x is defined*' stands for '$x \neq \uparrow$'. (E.g. $1/0$ is undefined, $1/0 = \uparrow$.)

- function symbols:[2]
  – (unary) 'absolute value' $|x|$, 'ceiling' $\lceil x \rceil$, 'weight' $\|x\|$. (For the meaning of $\|x\|$, see Theorem 1.26.)
  – (binary) $x+y, x-y, x.y, x/y, x\hat{\ }y$. (Axiom set 1.21 and Axiom 1.42.(4) assert that $x\hat{\ }n = x^n$ for hypernatural n, else undefined.)
  – for each $k \in \mathbb{N}$, k k-ary function symbols $\pi_{k,i}$ $(i = 1, \ldots, k)$. (Axiom schema 1.22 asserts that $\pi_{k,i}(\vec{x})$ are the projections of the k-tuple $\vec{x}$.)
  – for each quantifier-free formula φ with $m+1$ free variables, not involving min, an m-ary function symbol $\min_\varphi$. (For the meaning of which, see Theorems 1.35 and 1.41.)
  – for each triple $(k, \sigma(x_1, \ldots, x_m), \tau(x_1, \ldots, x_{m+2}))$ with $0 < k \in \mathbb{N}$, σ and τ terms not involving min, an $(m+1)$-ary function symbol $\mathrm{rec}^k_{\sigma\tau}$. (Axiom schema 1.31 asserts that this is the term obtained from σ and τ by recursion, after the model $f(0, \vec{x}) = \sigma(\vec{x})$, $f(n+1, \vec{x}) = \tau(f(n, \vec{x}), n, \vec{x})$, if terms are defined and do not weigh too much.)

1.8. Definition. If L is the language of ERNA, then L^{st}, the *standard* language of ERNA, is L without ω, ε or $\mathcal{I}$.

2.1.2. *The axioms of* ERNA.

1.9. Axiom set (Logic). *Axioms of first-order logic.*

1.10. Axiom set (Hypernaturals).

(1) 0 *is hypernatural*
(2) *if x is hypernatural, so is $x+1$*
(3) *if x is hypernatural, then $x \geq 0$*
(4) *ω is hypernatural.*

1.11. Definition. '*x is infinite*' stands for '$x \neq 0 \wedge 1/x \approx 0$'; '*$x$ is finite*' stands for 'x is not infinite'; '*x is natural*' stands for 'x is hypernatural and finite'.

1.12. Definition. A term or formula is called *internal* if it does not involve $\mathcal{I}$; if it does, it is called *external*.

1.13. Notation. The variables $n, m, k, l, \ldots$, both lower and upper-case, will represent hypernatural variables.

1.14. Axiom set (Infinitesimals).

(1) *if x and y are infinitesimal, so is $x+y$*
(2) *if x is infinitesimal and y is finite, xy is infinitesimal*
(3) *an infinitesimal is finite*
(4) *if x is infinitesimal and $|y| \leq x$, then y is infinitesimal*
(5) *if x and y are finite, so is $x+y$*

[2] We denote the values as computed in x or (x, y) according to the arity.

(6) *ε is infinitesimal*
(7) $\varepsilon = 1/\omega$.

1.15. Corollary. *1 is finite.*

Proof. If 1 is infinite, its inverse is infinitesimal, i.e. $1 \approx 0$. By axiom 1.14.(3), it would follow that 1 is finite, contradicting the assumption. □

1.16. Axiom set (Ordered field). *Axioms expressing that* ERNA*'s defined elements constitute an ordered field of characteristic zero with an absolute-value function. These quantifier-free axioms include*

- *if x is defined, then $x + 0 = 0 + x = x$*
- *if x is defined, then $x + (0 - x) = (0 - x) + x = 0$*
- *if x is defined and $x \neq 0$, then $x.(1/x) = (1/x).x = 1$.*

We write '$x < y$' instead of '$x \leq y \wedge \neg(x = y)$'.

1.17. Axiom (Archimedean). *If x is defined, $\big|\lceil x \rceil\big|$ is a hypernatural and $\lceil x \rceil - 1 < x \leq \lceil x \rceil$.*

1.18. Theorem. *If x is defined, then $\lceil x \rceil$ is the least integer $\geq x$.*

1.19. Theorem. *x is finite iff there is a natural n such that $|x| \leq n$.*

Proof. The statement is trivial for $x = 0$. If $x \neq 0$ is finite, so is $|x|$ because, assuming the opposite, $1/|x|$ would be infinitesimal and so would $1/x$ be by axiom 1.14.(4). By axiom 1.14.(5), the hypernatural $n = \lceil |x| \rceil < |x| + 1$ is then also finite. Conversely, let n be natural and $|x| \leq n$. If $1/|x|$ were infinitesimal, so would $1/n$ be by axiom 1.14.(4), and this contradicts the assumption that n is finite. □

1.20. Corollary. *$x \approx 0$ iff $|x| < 1/n$ for all natural $n \geq 1$.*

1.21. Axiom set (Power).

(1) *if x is defined, then $x\hat{}0 = 1$*
(2) *if x is defined and n is hypernatural, then $x\hat{}(n+1) = (x\hat{}n)\,x$.*

1.22. Axiom schema (Projection). *If $x_1, \ldots, x_n$ are defined, then we have $\pi_{n,i}(x_1, \ldots, x_n) = x_i$ for $i = 1, \ldots, n$.*

1.23. Axiom set (Weight).

(1) *if $\|x\|$ is defined, then $\|x\|$ is a nonzero hypernatural.*
(2) *if $|x| = m/n \leq 1$ (m and $n \neq 0$ hypernaturals), then $\|x\|$ is defined, $\|x\|.|x|$ is hypernatural and $\|x\| \leq n$*
(3) *if $|x| = m/n \geq 1$ (m and $n \neq 0$ hypernaturals), then $\|x\|$ is defined, $\|x\|/|x|$ is hypernatural and $\|x\| \leq m$.*

1.24. Definition. A *(hyper)rational* is of the form $\pm p/q$, with p and $q \neq 0$ (hyper)natural. We also use 'standard' instead of 'rational'.

1.25. Notation. $(\forall^{st} x)\varphi(x)$ stands for $(\forall x)(x \text{ is standard} \to \varphi(x))$ and $(\exists^{st} x)\varphi(x)$ for $(\exists x)(x \text{ is standard} \wedge \varphi(x))$.

1.26. Theorem.

(1) *If x is not a hyperrational, then $\|x\|$ is undefined.*
(2) *If $x = \pm p/q$ with p and $q \neq 0$ relatively prime hypernaturals, then*

$$\| \pm p/q\| = \max\{|p|, |q|\}.$$

1.27. Theorem.

(1) $\|0\| = 1$
(2) *if* $n \geq 1$ *is hypernatural,* $\|n\| = n$
(3) *if* $\|x\|$ *is defined, then* $\|1/x\| = \|x\|$ *and* $\| \lceil x \rceil \| \leq \|x\|$
(4) *if* $\|x\|$ *and* $\|y\|$ *are defined,* $\|x+y\|$, $\|x-y\|$, $\|xy\|$ *and* $\|x/y\|$ *are at most equal to* $(1+\|x\|)(1+\|y\|)$, *and* $\|x\hat{\ }y\|$ *is at most* $(1+\|x\|)\hat{\ }(1+\|y\|)$.

1.28. Notation. For any $0 < n \in \mathbb{N}$ we write $\|(x_1, \ldots, x_n)\|$ for the term $\max\{\|x_1\|, \ldots, \|x_n\|\}$.

1.29. Notation. For any $0 < n \in \mathbb{N}$ we write

$$2^x_n := \underbrace{2\,\hat{\ }\,(\ldots 2\,\hat{\ }\,(2\,\hat{\ }\,(2\,\hat{\ }\,x)))}_{n\ 2\text{'s}}.$$

1.30. Theorem. *If the term* $\tau(\vec{x})$ *is defined and does not involve* ω, rec *or* min, *then there exists a* $0 < k \in \mathbb{N}$ *such that*

$$\|\tau(\vec{x})\| \leq 2_k^{\|\vec{x}\|}.$$

1.31. Axiom schema (Recursion). *For any* $0 < k \in \mathbb{N}$ *and internal* σ, τ *not involving* min*:*

$$\mathrm{rec}^k_{\sigma\tau}(0,\vec{x}) = \begin{cases} \sigma(\vec{x}) & \textit{if this is defined, and has weight} \leq 2_k^{\|\vec{x}\|}, \\ \uparrow \textit{ if } \sigma(\vec{x}) = \uparrow, \\ 0 \textit{ otherwise.} \end{cases}$$

$$\mathrm{rec}^k_{\sigma\tau}(n+1,\vec{x}) = \begin{cases} \tau(\mathrm{rec}^k_{\sigma\tau}(n,\vec{x}),n,\vec{x}) & \textit{if defined, with weight} \leq 2_k^{\|\vec{x},n+1\|}, \\ \uparrow \textit{ if } \tau(\mathrm{rec}^k_{\sigma\tau}(n,\vec{x}),n,\vec{x}) = \uparrow, \\ 0 \textit{ otherwise.} \end{cases}$$

If the list $\vec{x}$ is empty, the above reduces to

$$\mathrm{rec}^k_{\sigma\tau}(0) = \sigma,$$

$$\mathrm{rec}^k_{\sigma\tau}(n+1) = \begin{cases} \tau(\mathrm{rec}^k_{\sigma\tau}(n),n) & \text{if defined, with weight} \leq 2_k^{n+1}, \\ \uparrow \text{ if } \tau(\mathrm{rec}^k_{\sigma\tau}(n),n) = \uparrow, \\ 0 \text{ otherwise.} \end{cases}$$

1.32. Corollary. *If* $\mathrm{rec}^k_{\sigma\tau}(n,\vec{x})$ *is defined, then* $\|\mathrm{rec}^k_{\sigma\tau}(n,\vec{x})\| \leq 2_k^{\|\vec{x},n\|}$.

We now adapt theorem 1.30 so as to allow more general terms.

1.33. Theorem.

(1) *If the term* $\tau_1(\vec{x})$ *is defined and does not involve* ω *or* min, *then there exists a* $0 < k \in \mathbb{N}$ *such that* $\|\tau_1(\vec{x})\| \leq 2_k^{\|\vec{x}\|}$.
(2) *If the term* $\tau_2(\vec{x})$ *is defined and does not involve* min, *then there exists a* $0 < k \in \mathbb{N}$ *such that* $\|\tau_2(\vec{x})\| \leq 2_k^{\|\vec{x},\omega\|}$.

Proof. For (1), we replace in $\tau_1(\vec{x})$ every term $\mathrm{rec}^k_{\sigma\tau}(n,\vec{y})$ by the corresponding term $2_k^{\|\vec{y},n\|}$. For the resulting term $\tau_1'(\vec{x})$ we have $\|\tau_1(\vec{x})\| \leq \|\tau_1'(\vec{x})\|$ by the preceding corollary. As the new term is defined and does not involve ω, min or rec, theorem 1.30 implies there is a $0 < k \in \mathbb{N}$ such that $\|\tau_1'(\vec{x})\| \leq 2_k^{\|\vec{x}\|}$. For (2), let $\tau_2'(\vec{x},m)$ be the term obtained by replacing, in $\tau_2(\vec{x})$, every occurrence of ω by m, and every occurrence of ε by $1/m$. By the previous item, there is a $0 < k \in \mathbb{N}$ such that $\|\tau_2'(\vec{x},m)\| \leq 2_k^{\|\vec{x},m\|}$. Hence $\|\tau_2(\vec{x})\| = \|\tau_2'(\vec{x},\omega)\| \leq 2_k^{\|\vec{x},\omega\|}$. □

1.34. Axiom schema (Internal minimum). *For every internal quantifier-free formula* $\varphi(y, \vec{x})$ *not involving* min *we have*

(1) $\min_\varphi(\vec{x})$ *is a hypernatural number*
(2) *if* $\min_\varphi(\vec{x}) > 0$, *then* $\varphi(\min_\varphi(\vec{x}), \vec{x})$
(3) *if* n *is a hypernatural and* $\varphi(n, \vec{x})$, *then* $\min_\varphi(\vec{x}) \leq n$ *and* $\varphi(\min_\varphi(\vec{x}), \vec{x})$.

1.35. Theorem. *If the internal quantifier-free formula* $\varphi(y, \vec{x})$ *does not involve* min, *and if there are hypernatural* n*'s such that* $\varphi(n, \vec{x})$, *then* $\min_\varphi(\vec{x})$ *is the least of these. If there are none,* $\min_\varphi(\vec{x}) = 0$.

1.36. Theorem (Hypernatural induction). *Let* $\varphi(n)$ *be an internal quantifier-free formula not involving* min, *such that*

(1) $\varphi(0)$
(2) $\varphi(n) \to \varphi(n+1)$.

Then $\varphi(n)$ *holds for all hypernatural* n.

Proof. Suppose, on the contrary, that there is a hypernatural n such that $\neg\varphi(n)$. By Theorem 1.35, there is a least hypernatural n_0 such that $\neg\varphi(n_0)$. By our assumption (1), $n_0 > 0$. Consequently, $\varphi(n_0 - 1)$ does hold. But then, by our assumption (2), so would $\varphi(n_0)$. This contradiction proves the theorem. □

1.37. Example. *If* $f(n)$ *is an internal function not involving* min *and such that* $0 < f(n) \leq \omega$ *for all* n, *then* $0 < \mathrm{rec}^1_{0f}(n) \leq \omega$ *for all* $n > 0$.

A more important application is hypernatural overflow and underflow in ERNA.

1.38. Theorem. *Let* $\varphi(n)$ *be an internal q.f. formula, not involving* min.

(1) *If* $\varphi(n)$ *holds for every natural* n, *it holds for all hypernatural* n *up to some infinite hypernatural* $\overline{n}$ ***(overflow)***.
(2) *If* $\varphi(n)$ *holds for every infinite hypernatural* n, *it holds for all hypernatural* n *from some natural* $\underline{n}$ *on* ***(underflow)***.

Proof. If $\varphi(n)$ holds for every hypernatural n, any $\overline{n}$ and $\underline{n}$ will do. If not, $n_0 = \min_{\neg\varphi}$ is the least hypernatural for which φ does not hold, and $n_1 = \min_{\neg\varphi'}$ is the least hypernatural for which $\varphi'(n) := \varphi(\omega - n)$ does not hold. By the assumption in (1), n_0 is infinite, and it follows that $\varphi(n)$ holds for every hypernatural $n \leq \overline{n} := n_0 - 1$. By the assumption in (2), $\omega - n_1$ is finite, and so is $\underline{n} := \omega - n_1 + 1$. For $\underline{n} \leq n \leq \omega$ we have $0 \leq \omega - n \leq n_1 - 1$, implying that $\varphi'(\omega - n) = \varphi(n)$ holds. For $n > \omega$, $\varphi(n)$ holds by assumption. Hence, $\underline{n}$ satisfies the requirements. □

This theorem allows us to prove Robinson's sequential lemma, (see [**48**, p. 150]), in ERNA.

1.39. Corollary. *Let* $f(n)$ *be an internal function not involving* min. *If* $f(n) \approx 0$ *for all* $n \in \mathbb{N}$, *then* $f(n) \approx 0$ *for all hypernatural* n *up to some infinite hypernatural* ω_1.

Proof. Apply overflow to the formula $|f(n)| < 1/n$. □

1.40. Axiom schema (External minimum). *For every (possibly external) q.f. formula* $\varphi(y, \vec{x})$ *not involving* min *or* ω *we have*

(1) $\min_\varphi(\vec{x})$ *is a hypernatural number*
(2) *if* $\min_\varphi(\vec{x}) > 0$, *then* $\varphi(\min_\varphi(\vec{x}), \vec{x})$

(3) *if n is a natural number, $\|\vec{x}\|$ is finite and $\varphi(n,\vec{x})$, then* $\min_\varphi(\vec{x}) \le n$ *and* $\varphi(\min_\varphi(\vec{x}),\vec{x})$.

1.41. Theorem. *Let $\varphi(n,\vec{x})$ be a (possibly external) quantifier-free formula not involving* min *or ω. If $\|\vec{x}\|$ is finite and if there are natural n's such that $\varphi(n,\vec{x})$, then* $\min_\varphi(\vec{x})$ *is the least of these. If there are none,* $\min_\varphi(\vec{x}) = 0$.

This theorem can be used to produce proofs by natural induction.

1.42. Axiom set ((Un)defined terms).

(1) $0, 1, \omega, \varepsilon$ *are defined*
(2) $|x|$, $\lceil x \rceil$, $\|x\|$ *are defined iff x is*
(3) $x + y$, $x - y$, xy *are defined iff x and y are; x/y is defined iff x and y are and $y \neq 0$*
(4) $x\hat{\ }y$ *is defined iff x and y are and y is hypernatural*
(5) $\pi_{k,i}(x_1,\dots,x_k)$ *is defined iff $x_1,\dots,x_k$ are*
(6) *if x is not a hypernatural,* $\mathrm{rec}^k_{\sigma\tau}(x,\vec{y})$ *is undefined*
(7) $\min_\varphi(x_1,\dots,x_k)$ *is defined iff $x_1,\dots,x_k$ are.*

1.43. Corollary. *In* ERNA*, 'defined' and 'hyperrational' mean the same.*

Proof. Let x be non-hyperrational. From theorem 1.26 we obtain that $\|x\|$ is undefined, and so is x by item (2) of the last axiom set. Hence, $\uparrow$ is the only non-hyperrational element in ERNA. $\square$

Note that ERNA has no 'standard-part' function st with the property that $\mathrm{st}(\varepsilon) = 0$ for $\varepsilon \approx 0$, which would allow for the unique decomposition of a finite number as the sum of a standard and an infinitesimal number, sometimes called the 'fundamental theorem of Nonstandard Analysis', [**39**]. Indeed, with such function st, ERNA would allow to construct the field of real numbers. As ERNA's consistency is proved in PRA, the latter would also allow to construct the real number field, something which is known to be impossible, [**33**]. Although the real number field is not available in ERNA, the rationals will turn out to be dense in the finite part of ERNA's field, see theorem 1.50. Moreover, theorems 1.70 and 1.73 show that the absence of the standard-part function in ERNA is not a great loss.

In [**49**], the consistency of ERNA is proved in PRA. Careful inspection shows that this proof also goes through in $I\Delta_0 + \text{superexp}$. The consistency of ERNA also follows from theorem 2.9. The choice for PRA as a 'background theory' is of course motivated by historical reasons. Also, since consistency is a Π_1-statement, it does not matter whether we use the quantifier-free 'strict finitist' version of PRA (see [**51**]) or the usual version which involves first-order logic. Indeed, a simple proof-theoretic argument shows that both prove the same Π_1-statements.

2.1.3. *Bootstrapping ERNA.* A bootstrapping process involves the step-by-step definition of certain basic functions, accompanied by proofs of their properties. We largely skip this rather tedious procedure for ERNA and only highlight the main results. For a full technical account of the bootstrapping process, the reader is referred to [**28**, §5] or Appendix A of this dissertation. For the rest of the chapter, we assume that the function f and the quantifier-free formulas φ, ψ do not involve min, $\approx$ or $\uparrow$.

1.44. THEOREM. *Let $f(n) \in L^{st}$ be an* ERNA*-term with weight at most 2_k^n for fixed $k \in \mathbb{N}$. Then $\sum_{n=0}^{m} f(n)$ and $\prod_{n=0}^{m} f(n)$ are* ERNA*-terms with weight at most 2_{k+2}^n.*

Thus, ERNA's functions are closed under sum and product. Once this has been established, it is easy to equip ERNA with pairing functions, used to reduce multivariable formulas to single-variable ones. This closure property also allows us to resolve bounded quantifiers. However, we first need the following theorem, interesting in its own right.

1.45. THEOREM. *For every internal quantifier-free formula $\varphi(\vec{x})$,* ERNA *has a function $T_\varphi(\vec{x})$ such that*

$$\varphi(\vec{x}) \text{ is true if and only if } T_\varphi(\vec{x}) = 1$$
$$\varphi(\vec{x}) \text{ is false if and only if } T_\varphi(\vec{x}) = 0.$$

1.46. COROLLARY. *For every internal quantifier-free formula $\varphi(n)$ and every hypernatural n_0, the formula $(\forall n \leq n_0)\varphi(n)$ is equivalent to $\prod_{n=0}^{n_0} T_\varphi(n) > 0$ and, likewise, $(\exists n \leq n_0)\varphi(n)$ is equivalent to $\sum_{n=0}^{n_0} T_\varphi(n) > 0$.*

Iterating and combining, we see that, as long as its quantifiers apply to bounded hypernatural variables, every internal formula not involving min or $\uparrow$ can be replaced by an equivalent quantifier-free one.

Essentially, the same result is also proved for the reduced Chuaqui and Suppes system NQA$^-$ in lemma 2.4 of [**42**].

Theorem 1.26 allows us to generalize the preceding corollary as follows.

1.47. COROLLARY. *For every internal quantifier-free formula $\varphi(x)$ not involving* min *or $\uparrow$ and every hypernatural n_0, the sentences $(\exists x)(\|x\| \leq n_0 \wedge \varphi(x))$ and $(\forall x)(\|x\| \leq n_0 \to \varphi(x))$ are equivalent to quantifier-free ones.*

Next, we consider a constructive version of theorem 1.38. Avoiding the use of $\min_\varphi$, it results in functions that can be used in recursion.

1.48. THEOREM. *Let $\varphi(n) \in \Delta_0$ be internal,*

(1) *If $\varphi(n)$ holds for every natural n, it holds for all hypernatural n up to some infinite hypernatural $\overline{n}$ **(overflow)**.*
(2) *If $\varphi(n)$ holds for every infinite hypernatural n, it holds for all hypernatural n from some natural $\underline{n}$ on **(underflow)**.*

Both numbers $\overline{n}$ and $\underline{n}$ are given by explicit ERNA*-formulas not involving* min.

This theorem has some immediate consequences.

1.49. COROLLARY.
Let φ be as in the theorem and assume $n_0 \in \mathbb{N}$.

(1) *If $\varphi(n)$ holds for every natural $n \geq n_0$, it holds for all hypernatural $n \geq n_0$ up to some infinite hypernatural $\overline{n}$, independent of n_0.*
(2) *If $\varphi(n_1, \ldots, n_k)$ holds for all natural $n_1, \ldots, n_k$, it holds for all hypernatural $n_1, \ldots, n_k$ up to some infinite hypernatural $\overline{n}$.*

In both cases the number $\overline{n}$ is given by an explicit ERNA*-formula.*

Analogous formulas hold for underflow. Overflow also allows us to prove that the rationals are dense in the finite hyperrationals, being ERNA's version of the 'fundamental theorem of Nonstandard Analysis'.

1.50. Theorem. *For every finite a and every natural n there is a rational b such that $|a-b|<\frac{1}{n}$.*

The following theorem is the dual of the previous.

1.51. Theorem. *In* ERNA*, there are hyperrationals of arbitrarily large weight between any two numbers.*

The following theorem shows that ERNA's functions are closed under the well-known bounded minimum.

1.52. Theorem. *Let φ be an internal Δ_0-formula. The bounded minimum*

$$(\mu n \leq M)\varphi(n,\vec{x}) := \begin{cases} \text{the least } n \leq M \text{ such that } \varphi(n,\vec{x}) & \text{if such exists,} \\ 0 & \text{otherwise,} \end{cases}$$

is definable in ERNA *using only sums and products.*

1.53. Theorem. *In* ERNA*, there are functions* 'max' *and* 'least' *which calculate the largest and the least entry from a list $(x_1,\ldots,x_k)$.*

1.54. Notation. We write $(\forall\omega)\varphi(\omega,\vec{x})$ for $(\forall n)(n \text{ is infinite} \to \varphi(n,\vec{x}))$. Likewise, $(\exists\omega)\varphi(\omega,\vec{x})$ means $(\exists n)(n \text{ is infinite} \wedge \varphi(n,\vec{x}))$.

The following theorem generalizes overflow to special external formulas.

1.55. Theorem. *Let ω_1 be infinite.*

(1) *If $f(n)$ is infinite for every $n\in\mathbb{N}$, it continues to be so for all hypernatural n up to some hypernatural number ω_2.*
(2) *If $(\forall^{st} n)(\exists\omega\leq\omega_1)\varphi(n,\omega)$, then there is an infinite hypernatural ω_3 such that $(\forall^{st} n)(\exists\omega\geq\omega_3)\varphi(n,\omega)$.*

2.2. ERNA and Transfer. In this section, we study several transfer principles from nonstandard mathematics in the context of ERNA. We are motivated by an interest in both metamathematical results *and* mathematical practice. Indeed, theorem 1.3 implies that transfer is essential for developing basic calculus in ERNA.

2.2.1. *ERNA and Universal Transfer.* In this paragraph, we add a transfer principle for universal sentences to ERNA and prove the consistency of the extended theory using a finite iteration of ERNA's consistency proof. We also show that this transfer principle is independent of ERNA.

1.56. Definition. If τ is an individual constant, the depth $d(\tau)$ is zero. For a term $\tau(x_1,\ldots,x_k)$ we put $d(\tau(x_1,\ldots,x_k)) = \max\{d(x_1),\ldots,d(x_k)\}+1$.

1.57. Axiom schema (Π_1-transfer). *For every quantifier-free formula $\varphi(n)$ from L^{st}, not involving* min*, we have*

$$\varphi(n+1) \vee \Big(0 < \min{}_{\neg\varphi} = \text{finite}\Big). \tag{1.1}$$

The above schema expresses in a quantifier-free way the basic transfer principle $(\forall^{st} n\geq 1)\varphi(n) \to (\forall n\geq 1)\varphi(n)$. After the consistency proof of ERNA + Π_1-TRANS$^-$, the reasons for the restrictions on φ will become apparent. We tacitly assume that *standard* parameters are allowed in φ in (1.1) and in all other (transfer) principles, unless explicitly stated otherwise. We use Π_1-TRANS to denote the previous axiom schema and we use Π_1-TRANS$^-$ to denote the parameter-free version of Π_1-TRANS. By theorems 1.3 and 1.130, the schemas Π_1-TRANS and Π_1-TRANS$^-$ play an important role in mathematical practice.

Before going into the consistency of ERNA + Π_1-TRANS$^-$, let us briefly review the consistency proof of ERNA. In view of Herbrand's theorem, we have to prove that any finite set T of instantiated axioms of ERNA is consistent. This we do by means of a mapping val. It maps all terms in T to functions of rationals and all relations in T to relations between rationals, in such a way that all the axioms in T receive the predicate 'true'. When this is achieved, T has a model.

The construction of val requires D steps, where D is the maximal depth of the finitely many terms occurring in T.

Three rational numbers $0 < a_0 < b_0 < c_0$ being chosen, ERNA's terms of zero depth are interpreted as $\mathrm{val}(0) = 0$, $\mathrm{val}(1) = 1$, $\mathrm{val}(\omega) = b_0$ and $\mathrm{val}(\varepsilon) = 1/b_0$.

After a finite number D of inductive steps, each one allowing terms of greater depth, all terms in T have been interpreted in such a way that $|\mathrm{val}(\tau)|$ belongs to $[0, a_D]$, $[b_D, c_D]$ or $[1/c_D, 1/b_D]$, according to τ being finite, infinite or infinitesimal. Finally $\mathrm{val}(x \approx 0)$ is defined by $|\mathrm{val}(x)| \leq 1/b_D$. Thus, all of ERNA's relations and terms have been given an interpretation. All that is left, is to check that all axioms in T receive the predicate 'true' under this interpretation. For this rather technical verification we refer to [**49**].

By theorem 1.33 there is a $0 < B \in \mathbb{N}$ such that for every term (of which there are only finitely many) $f(\vec{x})$ occurring in T, not involving min, we have

$$||f(\vec{x})|| \leq 2_B^{||\vec{x}||}. \tag{1.2}$$

Note that ω, which is allowed to occur, has been replaced with an extra free variable as in [**49**].

Then define

$$f_0(x) = 2_B^x \text{ and } f_{n+1}(x) = f_n^t(x) = \underbrace{f_n(f_n(\ldots(f_n(x))))}_{t\ f_n\text{'s}}, \tag{1.3}$$

$$a_0 = 1, b_0 = f_{D+1}(a_0), c_0 = b_0, d_0 = f_{D+1}(c_0) \tag{1.4}$$

and

$$a_{i+1} = f_{D-i}^j(a_i), b_i = f_{D-i}^{j+1}(a_i), c_{i+1} = f_{D-i}^l(c_i), d_{i+1} = f_{D-i}^{l+1}(c_i).$$

The numbers t, j and l are determined by the terms in the set T, their depths and the bounds on their weight; see [**49**] for details. Note that if we increase B to $B' > B$ and use $f_0'(x) = 2_{B'}^x$, the same D-step process as above would still yield a valid val$'$ for T. The same is true for increasing e.g. c_0. Also, $\mathrm{val}(\varphi(\vec{x})) = \varphi(\mathrm{val}(\vec{x}))$ for every quantifier-free formula φ of L^{st} not involving min; see [**49**] for details.

1.58. Theorem. ERNA + Π_1-TRANS$^-$ *is consistent and this consistency can be proved by a finite iteration of* ERNA*'s consistency proof.*

Proof. Let T be any finite set of instantiated axioms of ERNA+Π_1-TRANS$^-$. Let D be the maximum depth of the terms in T. Let $\varphi_1(n), \ldots, \varphi_N(n)$ be the quantifier-free formulas from L^{st} whose Π_1-transfer axiom (1.1) occurs in T. Leaving out these axioms from T, we are left with a finite set T' of instantiated ERNA-axioms. Let val be its interpretation into the rationals as sketched above. If we have

$$(\forall i \in \{1,\ldots,N\})\Big((\exists m \le a_D)\neg\varphi_i(m) \vee (\forall n \in [0,a_D] \cup [b_D,c_D])\varphi_i(n)\Big), \quad (1.5)$$

recalling that val maps finite numbers into $[0_D, a_D]$, we see that val provides a true interpretation of the whole of T, not just T'. On the other hand, assume there is an exceptional $\varphi' := \varphi_i$ for which

$$(\forall m \le a_D)\varphi'(m) \wedge (\exists n \in [0,a_D] \cup [b_D,c_D])\neg\varphi'(n). \quad (1.6)$$

Note that this implies $(\exists n \in [b_D,c_D])\neg\varphi'(n)$. Now choose a natural $B' > B$ such that $2^1_{B'} > c_D$, redefine $f_0(x)$ as $2^x_{B'}$ and construct an interpretation val$'$ in the same way as before. This val$'$ continues to make the axioms in T' true and does the same with the axiom

$$\varphi'(n+1) \vee (0 < \min_{\neg\varphi'} = \text{finite}). \quad (1.7)$$

Indeed, if a hypernatural n with $\text{val}(n) \in [b_D, c_D]$ makes φ' false, it is interpreted by val$'$ as a finite number because $n \le c_D \le a'_D$ by our choice of B'. Then the sentence $(\exists n \le a'_D)\neg\varphi'(n)$ is true; hence, $(0 < \min_{\neg\varphi'} = \text{finite})$ is true under val$'$ and so is the whole of (1.7).

Define T'' as T' plus all instances of (1.7) occurring in T. If there is another exceptional φ_i such that (1.6) holds, repeat this process. Note that if we increase B' to $B'' > B'$, redefine $f_0(x)$ as $2^x_{B''}$ and construct val$''$, the latter still makes the axioms of T' true, but the axioms of T'' as well, since $a'_D \le a''_D$ and hence (1.7) is true under val$''$ for the same reason as for val$'$.

This process, repeated, will certainly halt: either the list $\{1,\ldots,N\}$ becomes exhausted or, at some earlier stage, a valid interpretation is found for T. Note that this consistency proof, requiring at most ND steps, is a finite iteration of ERNA's, which requires at most D steps. □

The restrictions on the formulas φ admitted in (1.1) are imposed by our consistency proof. Neither $\approx$ nor ω can occur, because in ERNA's consistency proof, ω is interpreted as b_0 and '$x \approx 0$' as '$|x| \le 1/b_D$', both of which depend on B. By our changing B into $B' > B$, formulas like (1.7) could loose their 'true' interpretation from one step to the next. The exclusion of min has, of course, a different reason: $\min_\varphi$ is only allowed in ERNA when φ does not rely on min. Finally, theorem 1.75 shows that there is an interpretation of $I\Sigma_1$ in ERNA+Π_1-TRANS. Hence, we need to restrict Π_1-TRANS to the parameter-free schema Π_1-TRANS$^-$ to guarantee a finitistic consistency proof and to avoid contradicting Gödel's second incompleteness theorem.

Note that Parsons' theorem (see [**8**]) allows a shortcut in our consistency proof. To this end, we apply a certain algorithm $\mathcal{A}$ to our set of instantiated axioms T. The algorithm is as follows: construct val for T' and check whether it makes all the axioms in $T \setminus T'$ true; if so, return B; if not, add 1 to B and repeat as long as it takes to make all the axioms in $T \setminus T'$ true. The worst case is that every φ_i has a counterexample n_i, compelling the algorithm to possibly run until B is so large that a_D surpasses every $\min_{\neg\varphi_i}$. The 'while'-loop seems to carry this proof outside PRA, but this is not the case. By Parsons' theorem, if $I\Sigma_1$ proves that for every x there is a unique value $f(x)$, then the function $f(x)$ is primitive recursive. Equivalently, if $I\Sigma_1$ proves that an algorithm (possibly containing while-loops) halts for every input, then the algorithm is actually primitive recursive. The latter is the case for our algorithm $\mathcal{A}$: it only has to run until $a_D > \max_{1\leq i\leq N} \min_{\neg\varphi_i}$, which minorant is a term of $I\Sigma_1$. Our direct approach, used above, avoids this advanced conservation result, at the cost of greater length, but with a better bound on the strength of the 'background theory'.

Using results from the bootstrapping process, we can easily prove the following multivariable form of transfer, not restricted to hypernatural variables.

1.59. Theorem (Multivariable Tranfer). *Assume $\varphi(x_1,\dots,x_k)$ is a Δ_0-formula of L^{st}. In* $\text{ERNA}+\Pi_1$-TRANS *the sentences*

$$(\forall^{st}x_1)\dots(\forall^{st}x_k)\varphi(x_1,\dots,x_k) \text{ and } (\forall x_1)\dots(\forall x_k)\varphi(x_1,\dots,x_k)$$

are equivalent, and likewise the sentences

$$(\exists x_1)\dots(\exists x_k)\varphi(x_1,\dots,x_k) \text{ and } (\exists^{st}x_1)\dots(\exists^{st}x_k)\varphi(x_1,\dots,x_k).$$

It is well-known that WKL is independent of RCA_0 (see [**46**]). As suggested by theorem 1.3, the Π_1-transfer principle corresponds to WKL and hence it is to be expected that the former is independent of ERNA. We have the following stronger theorem.

1.60. Theorem. *The schema* Π_1-TRANS$^-$ *is independent of* ERNA.

Proof. Let ERNA^{st} be the set of all ERNA-axioms not involving ω or $\approx$. By Gödel's second incompleteness theorem (see [**8**, Section 2.2.3]), applied to ERNA^{st}, there is a quantifier-free formula φ of L^{st} such that ERNA^{st} proves neither $(\forall n)\varphi(n)$ nor its negation $(\exists n)\neg\varphi(n)$. Hence, there is a model $\mathcal{M}$ of ERNA^{st} such that $\mathcal{M}\models(\exists n)\neg\varphi(n)$. Moreover, for every $k\in\mathbb{N}$, there is a model $\mathcal{M}_k$ of ERNA^{st} such that $\mathcal{M}_k\models \min_{\neg\varphi}>k$. If not, there would be some $k_0\in\mathbb{N}$ such that $\min_{\neg\varphi}\leq k_0$ holds in all models of ERNA^{st}. By completeness, ERNA^{st} would then prove that $\min_{\neg\varphi}\leq k_0$ and, depending on $(\forall n\leq k_0)\varphi(n)$ being true or false, it could prove either $(\forall n)\varphi(n)$ or $(\exists n)\neg\varphi(n)$.

Now let c be a new constant and consider the sentence $\Phi\equiv\neg\varphi(c)\wedge(\forall^{st}n)\varphi(n)$. We will prove the consistency of $\text{ERNA}+\Phi$, using Herbrand's theorem in the same way as we did for $\text{ERNA}+\Pi_1$-TRANS$^-$. Let T be any finite set of instantiated axioms of $\text{ERNA}+\Phi$. Let D be the maximum depth of the terms in T. Leaving out all instances of the axiom Φ from T, we are left with a finite set T' of instantiated ERNA-axioms. Let val be its interpretation into the rationals and assume the infinite numbers are interpreted into $[b_D,c_D]$. Finally, let $b_0\in\mathbb{N}$ be such that $\text{val}(\omega)=b_0$. By the first paragraph of this proof, there is a model $\mathcal{M}_{b_0}$ of ERNA^{st} in which $\min_{\neg\varphi}>b_0$ holds. Let m_0 be the interpretation of $\min_{\neg\varphi}$ in $\mathcal{M}_{b_0}$. If necessary, increase the parameter c_0 from (1.4) to make sure $m_0\in[b_D,c_D]$ (compare theorem 1.58). Let val$'$ be val with the increased parameter c_0. Then the interpretation val$'$ with $\mathcal{M}_{b_0}$ as domain is also a valid interpretation for T'. Finally, defining $\text{val}'(c)=m_0$, we give Φ a valid interpretation too. Hence, all of T has received a valid interpretation and, by Herbrand's theorem, there is a model of ERNA in which Π_1-TRANS$^-$ is false.

For a model of ERNA in which Π_1-TRANS$^-$ is true, see theorem 1.58. Consequently, the independence is established. □

Note that the techniques used to prove the consistency of $\text{ERNA}+\Pi_1$-TRANS$^-$ and $\text{ERNA}+\Phi$ are essentially one and the same, applied in different directions.

Indeed, in the consistency proof of $\text{ERNA} + \Pi_1\text{-TRANS}^-$ (see theorem 1.58), the counterexamples to (1.1) are pushed down into the finite numbers by increasing B. In the above proof, however, such a counterexample is pushed upwards, into the infinite numbers, in order to obtain a valid interpretation for Φ.

2.2.2. *ERNA and stronger Transfer.* In this paragraph, we study ERNA's version of the transfer principle for larger formulas classes. First, we consider transfer for Π_2-formulas, defined next.

1.61. Principle (Π_2-TRANS). *For every quantifier-free formula φ in L^{st}, we have*

$$(\forall^{st} n)(\exists^{st} m)\varphi(n,m) \leftrightarrow (\forall n)(\exists m)\varphi(n,m).$$

We postpone the consistency proof of $\text{ERNA} + \Pi_2\text{-TRANS}$ until later, as we need results from chapter II for this proof. For now, we point out corollary 2.72 which states that $\text{ERNA} + \Pi_2\text{-TRANS}$ is provably consistent.

Let $\text{ERNA}^{\emptyset}$ be ERNA without its minimization axioms. The following theorem does away with the external and internal minimum in the consistency proof of $\text{ERNA}+\Pi_2\text{-TRANS}$. The gain is considerable, because treating minimization takes up a large portion of the consistency proof of ERNA and NQA^- (see [**42**] and [**49**]).

1.62. Theorem ($\min_\varphi$-redundancy). *The theories* $\text{ERNA}^{\emptyset}+\Pi_2\text{-TRANS}$ *and* $\text{ERNA}+\Pi_2\text{-TRANS}$ *prove the same theorems.*

Proof. First we treat the external minimum schema. Assume $\varphi(n,\vec{x})$ as in axioms schema 1.40, i.e. quantifier-free and not involving ω or min. Fix a natural n. Let φ' be φ with all positive occurrences of $\tau_i(n,\vec{x}) \approx 0$ replaced with $(\forall^{st} n_i)(|\tau_i(n,\vec{x})| < 1/n_i)$, where n_i is a new variable not appearing in φ. Do the same for the negative occurrences, using new variables m_i. Bringing all quantifiers in $\varphi'(n,\vec{x})$ to the front, we obtain

$$(\exists^{st} m_1)\dots(\exists^{st} m_l)(\forall^{st} n_1)\dots(\forall^{st} n_k)\psi(n,\vec{x},\vec{n},\vec{m})$$

where ψ is quantifier-free and standard. By Σ_2-transfer, this is equivalent to

$$(\exists m_1)\dots(\exists m_l)(\forall n_1)\dots(\forall n_k)\psi(n,\vec{x},\vec{n},\vec{m}).$$

If we return the quantifiers to their original places, all external atomic formulas $\tau_i(n,\vec{x}) \approx 0$ have become $(\forall n_i)(|\tau_i(n,\vec{x})| < 1/n_i)$ or, equivalently, $\tau_i(n,\vec{x}) = 0$. If $\varphi''(n,\vec{x})$ is φ with all symbols $\approx$ replaced with $=$, we have proved that $\varphi''(n,\vec{x})$ is equivalent to $\varphi(n,\vec{x})$. By theorem 1.52, $\text{ERNA}^{\emptyset}$ has a function which calculates the least n such that $\varphi''(n,\vec{x})$, if such there are. This function replaces the external minimum operator $\min_\varphi$.

Now for the internal minimum schema. Assume $\varphi(n,\vec{x})$ as in schema 1.34, i.e. quantifier-free and not involving $\approx$ or min. Let $\varphi(n,\vec{x},m)$ be $\varphi(n,\vec{x})$ with all occurrences of ω replaced with the new variable m. By theorem 1.52, $\text{ERNA}^{\emptyset}$ has a function which, for every finite m, calculates the least $n \le \omega$ such that $\varphi(n,\vec{x},m)$, if such there are. Then the sentence

$$(\forall l \le \omega)\neg\varphi(l,\vec{x},m) \vee (\exists n \le \omega)\big(\varphi(n,\vec{x},m) \wedge (\forall k < n)\neg\varphi(k,\vec{x},m)\big)$$

is true for all natural m. Using Σ_1-transfer, we obtain

$$(\forall^{st} m)\Big((\forall^{st} l)\neg\varphi(l,\vec{x},m) \vee (\exists^{st} n)\big(\varphi(n,\vec{x},m) \wedge (\forall k < n)\neg\varphi(k,\vec{x},m)\big)\Big)$$

and Π_2-transfer implies

$$(\forall m)\Big((\forall l)\neg\varphi(l,\vec{x},m) \vee (\exists n)\big(\varphi(n,\vec{x},m) \wedge (\forall k < n)\neg\varphi(k,\vec{x},m)\big)\Big).$$

If we fix $m = \omega$, the skolemization of the resulting sentence is exactly the axiom of the internal minimum schema for $\varphi(n,\vec{x})$. Since a theory and its skolemization prove the same theorems, we are done. $\square$

Note that, in order to prove that standard terms are finite for finite input, one needs external induction, which is equivalent to external minimization. Hence, it is not possible to prove external minimization without transfer by arguing that, as φ does not involve ω, all terms appearing in φ are standard and hence not-infinitesimal, unless zero. Also, it is interesting to compare the first paragraph of the proof to the part of ERNA's consistency proof that deals with the external minimum ([**49**]).

Next, we consider the transfer principle for Π_3-formulas.

1.63. Principle (Π_3-TRANS). *For each quantifier-free formula* $\varphi \in L^{st}$

$$(\forall^{st} n)(\exists^{st} m)(\forall^{st} k)\varphi(n,m,k) \leftrightarrow (\forall n)(\exists m)(\forall k)\varphi(n,m,k). \tag{1.8}$$

1.64. Theorem. *The theory* ERNA + Π_3-TRANS *proves induction for internal* Σ_1*-formulas.*

Proof. Below, we prove the Σ_1-induction axioms with standard quantifiers $(\forall^{st} n)$ and $(\exists^{st} m)$ instead of $(\forall n)$ and $(\exists m)$. As the (parametrized) Σ_1-induction axioms are Π_3, the theorem then follows, by Π_3-transfer.

The theory ERNA has minimization axioms (axiom schema 1.34) for internal quantifier-free formulas. By theorem 1.35 and corollary 1.46, ERNA proves the induction axioms for internal Δ_0-formulas. In the same way, ERNA proves the induction formulas for internal Δ_0-formulas, but with standard quantifiers $(\forall^{st} n)$ and $(\exists^{st} m)$ instead of the unbounded quantifiers $(\forall n)$ and $(\exists m)$.

Let $\varphi(n,m,\vec{x})$ be a Δ_0-formula of L^{st}. Fix a standard $\vec{x}$ and assume

$$(\exists^{st} n)\varphi(n,0,\vec{x}) \text{ and } (\forall^{st} m)\big((\exists^{st} n)\varphi(n,m,\vec{x}) \to (\exists^{st} n)\varphi(n,m+1,\vec{x})\big). \tag{1.9}$$

First of all, $(\exists^{st} n)\varphi(n,0,\vec{x})$ implies $(\exists n \le \omega)\varphi(n,0,\vec{x})$. Also, the second part of (1.9) implies $(\exists n \le \omega)\varphi(n,m,\vec{x}) \to (\exists n \le \omega)\varphi(n,m+1,\vec{x})$ for all $m \in \mathbb{N}$. Indeed, if $m \in \mathbb{N}$, then, by Σ_1-transfer, $(\exists n \le \omega)\varphi(n,m,\vec{x})$ implies $(\exists^{st} n)\varphi(n,m,\vec{x})$, which implies $(\exists^{st} n)\varphi(n,m+1,\vec{x})$ by (1.9), yielding $(\exists n \le \omega)\varphi(n,m+1,\vec{x})$. Thus, the previous implies

$$(\exists n \le \omega)\varphi(n,0,\vec{x}) \text{ and } (\forall^{st} m)\big[(\exists n \le \omega)\varphi(n,m,\vec{x}) \to (\exists n \le \omega)\varphi(n,m+1,\vec{x})\big].$$

By Δ_0-induction, we obtain $(\forall^{st} m)(\exists n \le \omega)\varphi(n,m,\vec{x})$ and Σ_1-transfer implies that $(\forall^{st} m)(\exists^{st} n)\varphi(n,m,\vec{x})$ and we are done. $\square$

Let Π_4-TRANS be the transfer principle 1.63 generalised to Π_4-formulas. We have the following theorem.

1.65. Theorem. *The theory* NQA^+ + Π_4-TRANS *proves induction for internal* Σ_2*-formulas.*

PROOF. In the proof of the previous theorem, we showed, using Π_1-transfer, that Σ_1-formulas with standard quantifier $(\exists^{st} n)$ and standard parameters are equivalent to nonstandard bounded formulas. In the same way, one proves that a formula $(\exists^{st} n)(\forall^{st} m)\varphi(n,m)$, with quantifier-free $\varphi \in L^{st}$, is equivalent to the external quantifier-free formula

$$(\mu n \leq \omega)(\forall m \leq \omega)\varphi(n,m) \text{ is finite.}$$

Thus, the external minimum schema of NQA^+ implies the Σ_2-induction axiom schema with standard quantifiers $(\forall^{st} n)$ and $(\exists^{st} m)$ instead of $(\forall n)$ and $(\exists m)$. As the (parametrized) Σ_2-induction axioms are Π_4, the theorem then follows, by Π_4-transfer. □

Thus, Π_4-transfer is highly unsuitable for finitist reasoning. Indeed, by the theorem, Π_4-transfer makes ERNA at least as strong as $I\Sigma_2$, which proves the totality of the Ackermann function. It is well-known that this function is not primitive recursive, i.e. not definable in PRA.

2.2.3. ERNA *and Generalized Transfer.* In this paragraph, we expand the scope of ERNA's transfer principles, which, until now, was quite limited. Indeed, both Π_1 and Σ_2-transfer are limited to formulas of L^{st}. Hence, a formula cannot be transferred if it contains, for instance, ERNA's cosine $\sum_{n=0}^{\omega}(-1)^n \frac{x^{2n}}{(2n)!}$ or similar objects not definable in L^{st}. This is quite a limitation, especially for the development of basic analysis. In this paragraph, we overcome this problem by widening the scope of transfer so as to be applicable to objects like ERNA's cosine. For Σ_2 and special Π_1-formulas, this is not so difficult, but not so for general Π_1-formulas.

First we label some terms which, though not part of L^{st}, are 'nearly as good' as standard for the purpose of transfer.

1.66. DEFINITION. For $\tau(n,\vec{x}) \in L^{st}$, the term $\tau(\omega,\vec{x})$ is *near-standard* if

$$(\forall \vec{x})(\forall \omega', \omega'')(\tau(\omega',\vec{x}) \approx \tau(\omega'',\vec{x})). \tag{1.10}$$

An atomic inequality $\tau(\omega,\vec{x}) \leq \sigma(\omega,\vec{x})$ is called near-standard if both members are. Since $x = y$ is equivalent to $x \leq y \wedge x \geq y$, and $\mathcal{N}(x)$ to $\lceil x \rceil = |x|$, any internal formula $\varphi(\omega,\vec{x})$ can be assumed to consist entirely of atomic inequalities; it is called near-standard if it is made up of near-standard atomic inequalities.

Full transfer for near-standard formulas is impossible. Indeed, the implication $|x| < 1 \rightarrow \frac{1}{|x|} > 1 + \frac{1}{\omega}$ is true for all *standard* x, but false for $x = \frac{2\omega}{2\omega+1}$. However, the weaker implication $|x| < 1 \rightarrow \frac{1}{|x|} \gtrapprox 1 + \frac{1}{\omega}$ does hold for all x, and this is the idea behind generalized transfer, to be considered next. We need a few definitions, first 'positive' and 'negative' occurrence of subformulas (see [**8**, **10**]).

Intuitively speaking, an occurrence of a subformula B in A is positive (negative) if, after resolving the implications outside B and pushing all negations inward, but not inside B, there is no (one) negation in front of B. Thus, in

$$(\neg(B \rightarrow C) \wedge (D \rightarrow B)) \rightarrow \neg D,$$

all occurrences of B are negative, C has one positive occurrence and D occurs both positively and negatively. The formal definition is as follows.

1.67. DEFINITION. Given a formula A, an occurrence of a subformula B, and an occurrence of a logical connective α in A, we say that B is *negatively bound* by α if either α is a negation $\neg$ and B is in its scope, or α is an implication $\to$ and B is a subformula of the antecedent. The subformula B is said to *occur negatively* (positively) in A if B is negatively bound by an odd (even) number of connectives of A.

1.68. NOTATION. We write $a \ll b$ for $a \leq b \wedge a \not\approx b$ and $a \lessapprox b$ for $a \leq b \vee a \approx b$.

1.69. DEFINITION. Given a near-standard formula $\varphi(\vec{x})$, let $\overline{\varphi}(\vec{x})$ be the formula obtained by replacing every positive (negative) occurrence of a near-standard inequality $\leq$ with $\lessapprox$ ($\ll$).

1.70. THEOREM (Generalized Transfer). *Let $\varphi(x,y)$ and $\psi(x)$ be near-standard and quantifier-free. In* ERNA + Π_2-TRANS *we have that*

$$(\forall^{st} x)\psi(x) \to (\forall x)\overline{\psi}(x) \tag{1.11}$$

and

$$(\forall^{st} x)(\exists^{st} y)\varphi(x,y) \to (\forall x)(\exists y)\overline{\varphi}(x,y). \tag{1.12}$$

PROOF.
We will prove the Π_2-case (1.12); for a proof of the Π_1-case (1.13), omit one quantifier. Let φ be as in the theorem.

First, we treat the atomic case where φ is $\tau_1(\omega,x,y) \leq \tau_2(\omega,x,y)$. So assume $(\forall^{st} x)(\exists^{st} y)\varphi(x,y)$. We have to prove that $(\forall x)(\exists y)\tau_1(\omega,x,y) \lessapprox \tau_2(\omega,x,y)$. If not, $(\exists x)(\forall y)(\tau_1(\omega,x,y) \gg \tau_2(\omega,x,y))$. Since both τ_1 and τ_2 are near-standard, they vary infinitesimally if ω is replaced with another infinite hypernatural. This property implies $(\exists x)(\forall y)(\forall m \geq \omega)(\tau_1(m,x,y) \gg \tau_2(m,x,y))$, hence the *standard* sentence

$$(\exists x)(\exists n)(\forall y)(\forall m \geq n)(\tau_1(m,x,y) > \tau_2(m,x,y)).$$

Using Σ_2-transfer, we obtain

$$(\exists^{st} x)(\exists^{st} n)(\forall^{st} y)(\forall^{st} m \geq n)(\tau_1(m,x,y) > \tau_2(m,x,y)).$$

Let x_0 and n_0 be standard numbers such that $(\forall^{st} y)(\forall^{st} m \geq n_0)(\tau_1(m,x_0,y) > \tau_2(m,x_0,y))$. By Π_1-transfer, $(\forall y)(\forall m \geq n_0)(\tau_1(m,x_0,y) > \tau_2(m,x_0,y))$ and since n_0 is finite, we have $(\forall y)(\tau_1(\omega,x_0,y) > \tau_2(\omega,x_0,y))$. As x_0 is standard, $(\exists^{st} x)(\forall^{st} y)(\tau_1(\omega,x,y) > \tau_2(\omega,x,y))$, contradicting the assumption.

For the general case, assume to the contrary that $(\forall^{st} x)(\exists^{st} y)\varphi(x,y,\omega)$ and $(\exists x)(\forall y)\neg\overline{\varphi}(x,y,\omega)$, with all occurrences of ω as shown. First, we use induction on the number of connectives in φ to see that the only near-standard atomic subformulas in $\neg\overline{\varphi}$, if the negation has been pushed inwards, are formulas with $\ll$ or $\gg$. Hence, all atomic near-standard formulas in $\neg\overline{\varphi}$ are of the form $\tau(x,y,\omega) \ll \sigma(x,y,\omega)$. As τ and σ are near-standard, they vary infinitesimally if ω is replaced with another infinite hypernatural. Hence, all formulas $\tau(x,y,\omega) \ll \sigma(x,y,\omega)$ in $\neg\overline{\varphi}$ do not change truth value if ω is replaced with $m \geq \omega$. Thus, $\neg\overline{\varphi}(x,y,\omega)$ implies $\neg\overline{\varphi}(x,y,m)$, for all $m \geq \omega$. From $(\exists x)(\forall y)\neg\overline{\varphi}(x,y,\omega)$ there follows $(\exists x)(\forall y)(\forall m \geq \omega)\neg\overline{\varphi}(x,y,m)$, which implies $(\exists x)(\exists n)(\forall y)(\forall m \geq n)\neg\overline{\varphi}(x,y,m)$. In the same way as in the atomic case, we obtain the formula $(\exists^{st} x)(\forall^{st} y)\neg\varphi(x,y,\omega)$, which is a contradiction. $\square$

Let $\overline{\Pi}_2$-TRANS be the schema consisting of all formula (1.12) for near-standard quantifier-free φ. By the previous theorem, ERNA proves that this axiom schema is equivalent to Π_2-TRANS. Similar theorems exist for the formula classes Π_n for $n \geq 3$, and the proof would be essentially identical to the previous proof.

The near-standard condition (1.10) can be omitted in the special case we consider next.

1.71. Theorem (Generalized Transfer, special case). *Let $\psi(x)$ be a quantifier-free formula whose only nonstandard terms are finite and of the form $\tau(\omega)$, with τ internal. In* ERNA + Π_1-TRANS *we have that*

$$(\forall^{st} x)\psi(x) \to (\forall x)\overline{\psi}(x).$$

Proof. As in the previous proof, it suffices to consider the atomic case. Assume that $\tau_1(x)$ is standard and that $(\forall^{st} x)(\tau_1(x) \leq \tau(\omega))$, where $\tau(\omega)$ is finite. If $(\exists x)(\tau_1(x) \gg \tau(\omega))$, choose such an $x = x_0$. Then there is a rational number q such that $\tau_1(x_0) \geq q > \tau(\omega)$. From $(\exists x)(\tau_1(x) \geq q)$ we obtain by Σ_1-transfer that $(\exists^{st} x)(\tau_1(x) \geq q)$, hence $(\exists^{st} x)(\tau_1(x) > \tau(\omega))$. This contradicts the assumption. □

Thus, we can work freely in ERNA + Π_2-TRANS with functions such as ERNA's cosine. However, in ERNA + Π_1-TRANS, we can only work freely with constants such as $e := \sum_{n=0}^{\omega} \frac{1}{n!}$ and $\pi := 4\sum_{k=0}^{\omega} \frac{(-1)^k}{2k+1}$. Next, we show that Π_1-transfer also implies (1.13) and hence we can work freely with functions like ERNA's cosine in ERNA+Π_1-TRANS too. This is a key element in our study of Reverse Mathematics and requires considerable more technical effort than the proof of theorem 1.70.

First, consider the following transfer principle.

1.72. Principle ($\overline{\Pi}_1$-TRANS). *Let $\varphi(x)$ be near-standard and quantifier-free. Then,*

$$(\forall^{st} x)\varphi(x) \to (\forall x)\overline{\varphi}(x). \tag{1.13}$$

The previous principle is also called 'bar transfer'. When formulated in contrapositive form, bar transfer is called '$\overline{\Sigma}_1$-transfer'.

1.73. Theorem. *In* ERNA*,* Π_1-TRANS *and* $\overline{\Pi}_1$-TRANS *are equivalent.*

Proof. For a standard formula, we have $\overline{\varphi} \equiv \varphi$ and hence the schema $\overline{\Pi}_1$-TRANS clearly implies Π_1-TRANS. Now assume Π_1-TRANS, let φ be as in $\overline{\Pi}_1$-TRANS and let τ_1 and τ_2 be near-standard terms. We first prove the atomic case, i.e. where $\varphi(n)$ is $\tau_1(n,\omega) \leq \tau_2(n,\omega)$. So, assume that $\varphi(n)$ holds for all $n \in \mathbb{N}$, and consider the sentence

$$(\forall n)(\forall \omega', \omega'')(\tau_i(n,\omega') \approx \tau_i(n,\omega'')) \tag{1.14}$$

for $i = 1, 2$. This sentence expresses that τ_1 and τ_2 are near-standard. Also, it implies

$$(\forall^{st} k)(\forall n)(\forall \omega', \omega'')(|\tau_i(n,\omega') - \tau_i(n,\omega'')| < 1/k), \tag{1.15}$$

and also

$$(\forall^{st} k)(\forall \omega', \omega'')(\forall n \leq \omega_1)(|\tau_i(n,\omega') - \tau_i(n,\omega'')| < 1/k),$$

where ω_1 is a fixed infinite hypernatural number. By underflow, there follows

$$(\forall^{st} k)(\exists^{st} M)(\forall m, m' \geq M)(\forall n \leq \omega_1)(|\tau_i(n,m) - \tau_i(n,m')| < 1/k) \tag{1.16}$$

and Π_1-transfer implies

$$(\forall^{st}k)(\exists^{st}M)(\forall m, m' \geq M)(\forall n)(|\tau_i(n,m) - \tau_i(n,m')| < 1/k). \tag{1.17}$$

Now suppose there exists a number n_0 satisfying $\tau_1(n_0,\omega) \gg \tau_2(n_0,\omega)$ and assume $k_0 \in \mathbb{N}$ is such that $\tau_1(n_0,\omega) - \tau_2(n_0,\omega) > 1/k_0$. Then apply (1.17) for $k = 4k_0$ and obtain, for $i = 1, 2$, a number $M_i \in \mathbb{N}$ such that

$$(\forall n)(\forall m, m' \geq M_i)(|\tau_i(n,m) - \tau_i(n,m')| < 1/4k_0). \tag{1.18}$$

In particular, this implies

$$|\tau_i(n_0, M_i) - \tau_i(n_0,\omega)| < 1/4k_0,$$

for $i = 1, 2$. This formula, together with $\tau_1(n_0,\omega) - \tau_2(n_0,\omega) > 1/k_0$, implies

$$\tau_1(n_0, M_1) - \tau_2(n_0, M_2) > 1/2k_0, \tag{1.19}$$

yielding

$$(\exists n)(\tau_1(n, M_1) - \tau_2(n, M_2) > 1/2k_0).$$

By Σ_1-transfer, we obtain

$$(\exists^{st}n)(\tau_1(n, M_1) - \tau_2(n, M_2) > 1/2k_0).$$

By (1.18), this implies

$$(\exists^{st}n)(\tau_1(n,\omega) - \tau_2(n,\omega) > 0),$$

which contradicts our assumption $(\forall^{st}n)(\tau_1(n,\omega) \leq \tau_2(n,\omega))$.

For the general case, we use induction on the number of near-standard atomic formulas. We may assume that in $\overline{\varphi}$ each instance of $A \to B$ is replaced by $\neg A \vee B$ and that all negations have been pushed in front of the atomic formulas by using De Morgan's laws from left to right. By definition each instance of $a \ll b$ in $\overline{\varphi}$ occurs negatively and hence each instance of $a \ll b$ now occurs as $\neg(a \ll b)$. Thus, it can be replaced by $a \gtrapprox b$ and hence we may assume $\overline{\varphi}$ to be free of '$\ll$'.

In case only one near-standard atomic formula occurs in $\varphi(n)$, the latter has the form either $\tau_1(n,\omega) \leq \tau_2(n,\omega) \wedge \psi(n)$ or $\tau_1(n,\omega) \leq \tau_2(n,\omega) \vee \psi(n)$, with $\psi \in L^{st}$ quantifier-free. In the first case, consider $(\forall^{st}n)\varphi(n)$ and push the universal quantifier through the conjunction. Now apply regular Π_1-transfer to the second part of the conjunction and apply the atomic case treated above to the first part. Hence, there follows $(\forall n)\overline{\varphi}(n)$. For the second case, assume $(\forall^{st}n)\varphi(n)$ and suppose there is a number n_0 such that $\neg\overline{\varphi}(n_0)$, i.e. $\tau_1(n_0,\omega) \gg \tau_2(n_0,\omega) \wedge \neg\psi(n_0)$. Let k_0 be such that $\tau_1(n_0,\omega) - \tau_2(n_0,\omega) > 1/k_0$. In exactly the same way as in the atomic case above, we obtain (1.18) and (1.19). As there also holds $\neg\psi(n_0)$, (1.19) implies

$$(\exists n)\big[\tau_1(n, M_1) - \tau_2(n, M_2) > 1/2k_0 \wedge \neg\psi(n)\big].$$

The previous formula is standard and hence, by Σ_1-transfer, there follows

$$(\exists^{st}n)\big[\tau_1(n, M_1) - \tau_2(n, M_2) > 1/2k_0 \wedge \neg\psi(n)\big].$$

By (1.18), there follows

$$(\exists^{st}n)\big[\tau_1(n,\omega) - \tau_2(n,\omega) > 0 \wedge \neg\psi(n)\big].$$

This contradicts our assumption that $\varphi(n)$ holds for all $n \in \mathbb{N}$ and this case is done.

Now assume we have established the case for $m \geq 1$ occurrences of near-standard atomic formulas. We now prove bar transfer for formulas $\varphi(n)$ with $m + 1$ occurrences of near-standard atomic formulas. Again, the formula $\varphi(n)$ has the form

$\tau_1(n,\omega) \leq \tau_2(n,\omega) \wedge \psi(n)$ or $\tau_1(n,\omega) \leq \tau_2(n,\omega) \vee \psi(n)$, where ψ only has m occurrences of near-standard atomic formulas. The first case is treated in the same way as in the case for $m = 1$, with the exception that the induction hypothesis is invoked to apply bar transfer to $(\forall^{st} n)\psi(n)$. For the second case, assume $(\forall^{st} n)\varphi(n)$ and suppose there is a number n_0 such that $\neg\overline{\varphi}(n_0)$, i.e. $\tau_1(n_0,\omega) \gg \tau_2(n_0,\omega) \wedge \neg\overline{\psi}(n_0)$. Let k_0 be such that $\tau_1(n_0,\omega) - \tau_2(n_0,\omega) > 1/k_0$. In the same way as before, we obtain (1.18) and (1.19). As there also holds $\neg\overline{\psi}(n_0)$, (1.19) implies

$$(\exists n)\big[\tau_1(n, M_1) - \tau_2(n, M_2) > 1/2k_0 \wedge \neg\overline{\psi}(n)\big].$$

The previous formula only involves m occurrences of atomic near-standard formulas and hence the induction hypothesis applies to it. By $\overline{\Sigma}_1$-transfer, there follows

$$(\exists^{st} n)\big[\tau_1(n, M_1) - \tau_2(n, M_2) > 1/2k_0 \wedge \neg\psi(n)\big]. \tag{1.20}$$

By (1.18), there follows

$$(\exists^{st} n)\big[\tau_1(n,\omega) - \tau_2(n,\omega) > 0 \wedge \neg\psi(n)\big].$$

This contradicts our assumption that $\varphi(n)$ holds for all $n \in \mathbb{N}$ and we are done. □

We point out that without theorem 1.73 all items listed in theorem 1.3 would be limited to *standard* functions. This would exclude most functions of basic analysis, like ERNA's cosine and exponential. Furthermore, we note that theorems 1.70 and 1.73 show that the transfer principle of Nonstandard Analysis is also robust in the sense discussed earlier.

2.3. ERNA and the Chuaqui and Suppes system. The theory ERNA is based on an earlier system by Chuaqui and Suppes (see [**11**]). Recently, Rössler and Jeřábek weakened ERNA's predecessor, the Chuaqui and Suppes system NQA^+, into NQA^- by introducing a different axiom schema for external minimization ([**42**]). They also showed that NQA^- is more suitable than NQA^+ for finitistic reasoning in the sense of Tait ([**51**]). We also refer to $\mathrm{NQA}^\emptyset$, which is NQA^+ without minimization axioms.

Most (all) of our ERNA theorems can be proved in NQA^- (NQA^+) without much adaptation; for an example, see theorem A.4. The converse, of course, is not true. While ERNA and NQA^+ can prove that a standard term $\tau(\vec{x})$ has standard values for standard $\vec{x}$, NQA^-, lacking full external induction, could not.

Our consistency proof of $\mathrm{ERNA} + \Pi_1$-TRANS is a finite iteration of ERNA's. Likewise, that of $\mathrm{NQA}^\pm$+Π_1-TRANS would be a finite iteration of that for $\mathrm{NQA}^\pm$. Also, all theorems of $\mathrm{ERNA} + \Pi_1$-TRANS could be proved in NQA^++Π_1-TRANS and most would also in NQA^-+Π_1-TRANS if the transfer axiom is adapted accordingly. But transfer is too strong for finitism in the sense of Tait. This is evident from the next theorem, to be compared to lemma 4.2 in [**42**], from which we also adopt the notations.

1.74. Theorem. *The theory* WKL_0 *is interpretable in* $\mathrm{NQA}^\emptyset$+O-MIN +Π_1-TRANS *and in* $\mathrm{ERNA} + \Pi_1$-TRANS.

Proof. In [**42**], the interpretation of $I\Sigma_1$ in NQA^+ is based on replacing all arithmetical Σ_1-formulas with quantifications relativized to $\mathbb{FN}(x)$, which are in turn replaced by external open formulas, provided by lemma 4.2 of [**42**]. If this has been done, the Σ_1-induction axioms of $I\Sigma_1$ are interpreted as instances of external open induction, which are implied by the schema O-MIN^{st} of NQA^+.

For the interpretation of $I\Sigma_1$ in $\mathrm{NQA}^{\emptyset}$+O-MIN+Π_1-TRANS, we start from the same interpretation of arithmetical Σ_1-formulas as quantifications relativized to $\mathbb{FN}(x)$. Lemma 4.2 in [**42**] contains the $\mathrm{NQA}^{\emptyset}$-term

$$m_{\varphi,\nu_0}(\vec{x}) := (\mu y \leq \nu_0(t_{\varphi(y,\vec{x})}(y,\vec{x}) = 1)),$$

comparable to ERNA's bounded minimum. Now $\varphi(m_{\varphi,\nu_0}(\vec{x}),\vec{x})$ implies the formula $(\exists y)(\mathbb{N}(y) \wedge \varphi(y,\vec{x}))$ and from the latter we obtain $(\exists y)(\mathbb{FN}(y) \wedge \varphi(y,\vec{x}))$, as Σ_1-transfer is contained in $\mathrm{NQA}^{\emptyset}$+O-MIN+Π_1-TRANS. Since all standard numbers are smaller than ν_0, the formula $(\exists y)(\mathbb{FN}(y) \wedge \varphi(y,\vec{x}))$ implies the formula $\varphi(m_{\varphi,\nu_0}(\vec{x}),\vec{x})$ by the definition of m_{φ,ν_0}. Thus, $\mathrm{NQA}^{\emptyset}$+O-MIN+Π_1-TRANS proves the equivalence

$$(\exists y)(\mathbb{FN}(y) \wedge \varphi(y,\vec{x})) \leftrightarrow \varphi(m_{\varphi,\nu_0}(\vec{x}),\vec{x}).$$

This equivalence implies that, once all arithmetical Σ_1-formulas have been replaced with quantifications relativized to $\mathbb{FN}(x)$, the interpreted Σ_1-induction axioms of $I\Sigma_1$ are equivalent to instances of internal open induction and hence follow from O-MIN. In section 4.3 of [**42**] the interpretation of $I\Sigma_1$ in NQA^+ is extended to an interpretation of WKL_0 in NQA^+. Exactly the same technique can be applied here to obtain an interpretation of WKL_0 in $\mathrm{NQA}^{\emptyset}$+O-MIN +Π_1-TRANS. In exactly the same way, the theorem follows for ERNA + Π_1-TRANS. $\square$

As a generalization, the following theorem shows that even stronger theories like $I\Sigma_2$ and $B\Sigma_2$ can be interpreted in ERNA and NQA^+ plus transfer.

1.75. Theorem.

(1) *The theory* $I\Sigma_2$ *can be interpreted in* $\mathrm{NQA}^+ + \Pi_1$-TRANS.
(2) *The theory* $B\Sigma_2$ *can be interpreted in* ERNA $+ \Pi_1$-TRANS.

Proof. For the notations '$\mathbb{FN}(n)$', '$\mu m \leq \nu_0$' and 'O-MIN$^{\mathrm{st}}$', we again refer to [**42**]. Additionally we assume $n, m, k, l, \dots$ to be *hypernatural* variables, i.e. satisfying the predicate $\mathbb{N}$ of $\mathrm{NQA}^{\emptyset}$. To interpret $I\Sigma_2$ in $\mathrm{NQA}^+ + \Pi_1$-TRANS, we start from the interpretation of arithmetical Σ_2-formulas as quantifications relativized to $\mathbb{FN}(n)$. From [**42**, Lemma 2.4], it follows that the $\mathrm{NQA}^{\emptyset}$-term

$$m_{\varphi,\nu_0}(\vec{n}) := (\mu m \leq \nu_0(t_{(\forall k \leq \nu_0)\varphi(m,k,\vec{n})}(m,\vec{n}) = 1))$$

is definable in NQA^+. Now $\mathbb{FN}(m_{\varphi,\nu_0}(\vec{n}))$ implies the formula $(\exists m)(\mathbb{FN}(m) \wedge (\forall k \leq \nu_o)\varphi(m,k,\vec{n}))$, hence $(\exists m)\big(\mathbb{FN}(m) \wedge (\forall k)(\mathbb{FN}(k) \to \varphi(m,k,\vec{n})\big)$. On the other hand, if m_0 is such that $\mathbb{FN}(m_0) \wedge (\forall k)(\mathbb{FN}(k) \to \varphi(m_0,k,\vec{n})$, Π_1-transfer applied to $(\forall k)(\mathbb{FN}(k) \to \varphi(m_0,k,\vec{n}))$ implies $(\forall k)(\mathbb{N}(k) \to \varphi(m_0,k,\vec{n}))$, hence certainly $(\forall k \leq \nu_0)\varphi(m_0,k,\vec{n})$. Now $m_{\varphi,\nu_0}(\vec{n})$ is standard; in fact it is at most m_0, because it is the least of the m satisfying $(\forall k \leq \nu_0)\varphi(m,k,\vec{n})$. Thus, $\mathrm{NQA}^+ + \Pi_1$-TRANS proves the equivalence

$$(\exists m)\big(\mathbb{FN}(m) \wedge (\forall k)(\mathbb{FN}(k) \to \varphi(m,k,\vec{n})\big) \leftrightarrow \mathbb{FN}(m_{\varphi,\nu_0}(\vec{n})). \tag{1.21}$$

It follows that, all arithmetical Σ_2-formulas being replaced with quantifications relativized to $\mathbb{FN}(x)$, the interpreted Σ_2-induction axioms of $I\Sigma_2$ are equivalent to instances of external open induction. Hence, they follow from O-MIN$^{\mathrm{st}}$.

For the second, we also interpret the quantifiers $(\exists n)$ and $(\forall m)$, occurring in formulas of $B\Sigma_2$, as $(\exists^{st} n)$ and $(\forall^{st} m)$, respectively, in ERNA $+ \Pi_1$-TRANS. Fix $k_0 \in \mathbb{N}$

and let $\varphi(k,l)$ be the Σ_2 sentence $(\exists n)(\forall m)\varphi_0(n,m,k,l)$ with φ_0 quantifier-free. Then the interpretation of the antecedent of the REPL-axiom of $B\Sigma_2$ for φ is

$$(\forall k \leq k_0)(\exists^{st} l)(\exists^{st} n)(\forall^{st} m)\varphi_0(n,m,k,l).$$

Using Π_1-transfer for suitable $k,l,n \in \mathbb{N}$, we obtain

$$(\forall k \leq k_0)(\exists^{st} l)(\exists^{st} n)(\forall m)\varphi_0(n,m,k,l),$$

hence certainly

$$(\forall k \leq k_0)(\exists^{st} l)(\exists^{st} n)(\forall m \leq \omega)\varphi_0(n,m,k,l).$$

Using a binary pairing function, we reduce $(\exists^{st} l)$ and $(\exists^{st} n)$ to a single quantifier $(\exists^{st} N)$. By theorem 1.52, ERNA$^{\emptyset}$ has an internal function $f(k)$ which calculates the least of these. Defining $l_0 = \sum_{k=0}^{k_0} f(k)$, we find

$$(\forall k \leq k_0)(\exists l \leq l_0)(\exists n \leq l_0)(\forall m \leq \omega)\varphi_0(n,m,k,l),$$

which yields

$$(\exists^{st} l_0)(\forall k \leq k_0)(\exists l \leq l_0)(\exists^{st} n)(\forall^{st} m)\varphi_0(n,m,k,l),$$

i.e. the consequent of the interpretation of the REPL-axiom of φ. □

The previous theorem, together with theorem 1.3, shows that many theorems of ordinary mathematics go beyond PRA. However, these theorems are still part of 'finitistic reductionism' (a partial realization of Hilbert's program, see [**46**]) as a nonstandard version of PRA, extended with Π_2-transfer, is still conservative over PRA ([**1**]).

3. Mathematics in ERNA

3.1. Mathematics without Transfer. In this section, we obtain some well-known mathematical theorems in ERNA, without using the transfer principle. The 'running theme' is that ERNA can prove many theorems of ordinary mathematics, as long as we allow an infinitesimal error. This theme is best expressed in theorems 1.77, 1.81, 1.94 and 1.96.

We assume, for the rest of this chapter, that a and b are finite numbers such that $a \not\approx b$ and that $f(x)$ is an internal function, not involving min and everywhere defined.

3.1.1. *Continuity.* First, we introduce the notion of (nonstandard) continuity in ERNA and prove some fundamental results.

1.76. Definition. A function $f(x)$ is 'continuous over $[a,b]$' if

$$(\forall x,y \in [a,b])(x \approx y \to f(x) \approx f(y)). \tag{1.22}$$

The attentive reader has noted that we work with the nonstandard version of uniform continuity. There are two reasons for this. First of all, if we limit the variable x in (1.22) to $\mathbb{Q}$, the function $\frac{1}{x^2-2}$ satisfies the resulting formula, although this function is unbounded in the interval $[-2,2]$. Similarly, the function $g(x)$, defined as 1 if $x^2 < 2$ and 0 if $x^2 \geq 2$, satisfies (1.22) with x limited to $\mathbb{Q}$, but $g(x)$ has a jump in its graph. The same holds for the pointwise ε-δ continiuty and thus both are not suitable for our purposes. Second, in light of theorem 1.3, the ε-δ definition of uniform continuity is closely related to Π_1-transfer. In the absence of the latter principle, we are left with (1.22).

1.77. THEOREM (Weierstraß extremum theorem). *If f is continuous over $[a,b]$, there is a number $c \in [a,b]$ such that for all $x \in [a,b]$, we have $f(x) \lessapprox f(c)$.*

PROOF. Let a, b, f be as stated. The points $x_n = a + \frac{n(b-a)}{\omega}$, for hypernatural $1 \le n \le \omega - 1$, partition the interval $[a,b]$ in infinitesimal subintervals $[x_n, x_{n+1}]$. Every $x \in [a,b]$ is in one of these intervals, hence infinitely close to both of its end points. As f is continuous over $[a,b]$, we have $f(x) \approx f(x_n)$ for $x \in [x_n, x_{n+1}]$. By theorem 1.53, ERNA has an explicit maximum operator, which allows to define

$$M := \max_{0 \le n \le \omega - 1} f(x_n). \tag{1.23}$$

Hence, $f(x) \lessapprox M$ for all $x \in [a,b]$. □

1.78. COROLLARY. *If f is continuous on $[a,b]$ and finite in at least one point, then it is finitely bounded on $[a,b]$.*

PROOF. Let a, b, f be as stated. Denote by $\varphi(n)$ the formula

$$(\forall x, y \in [a,b])(|x-y| \le 1/n \wedge \|x,y\| \le \omega \rightarrow |f(x) - f(y)| < 1). \tag{1.24}$$

As f is continuous, $\varphi(n)$ holds for all infinite n. By corollary 1.47, (1.24) may be treated as quantifier-free. Underflow yields that it holds from some finite n_0 on. Assume $f(x_0)$ is finite in $x_0 \in [a,b]$. Partitioning $[a,b]$ with points $1/\omega$ apart shows that we may assume $\|x_0\| \le \omega$. Then $\varphi(n_0)$ implies that $f(c)$ given by the theorem deviates at most $n_0\lceil b-a \rceil$ from $f(x_0)$. □

1.79. COROLLARY. *If f is near-standard and cont. on $[a,b]$, it is finitely bounded there.*

PROOF. From (1.10), we can derive (1.16) for $f(x,\omega)$ instead of $\tau_i(n,\omega)$. Thus, f is finitely close to a standard term in at least one point. By theorem 1.33, this standard term is finite and hence f is finite in at least one point. □

1.80. THEOREM (Intermediate value theorem). *If f is continuous on $[a,b]$, and $f(a) \le y_0 \le f(b)$, then there is an $x_0 \in [a,b]$ such that $f(x_0) \approx y_0$.*

PROOF. Let a, b, y_0 and f be stated. The points $x_n = a + \frac{n(b-a)}{\omega}$, for hypernatural $1 \le n \le \omega - 1$, partition the interval $[a,b]$ in infinitesimal subintervals $[x_n, x_{n+1}]$. Similarly, the points $f(x_n)$ partition the interval $[f(a), f(b)]$ in subintervals. As f is continuous, the intervals $[f(x_n), f(x_{n+1})]$ are also infinitesimal. Using ERNA's explicit 'least' operator (see theorem 1.53), we find an $N \le \omega$ such that $|y_0 - f(x_N)|$ is minimal. Hence, we have $y_0 \in [f(x_N), f(x_{N+1})]$ or $y_0 \in [f(x_{N-1}), f(x_N)]$. In either case, $x_0 = x_N$ satisfies the requirements. □

Note that if there are rational x_1 and y_1 such that $x_0 \approx x_1$ and $y_0 \approx y_1$, then $y_0 \approx f(x_0)$ implies $y_1 = f(x_1)$, if f is continuous. Most numbers in ERNA, however, do not have a standard number infinitely close.

1.81. COROLLARY (Brouwer fixed point theorem). *If $f : [0,1] \to [0,1]$ is continuous, then there is an $x_0 \in [0,1]$ such that $f(x_0) \approx x_0$.*

PROOF. Let f be as stated. If $f(0) \approx 0$ or $f(1) \approx 1$, we are done. Otherwise, $f(1) - 1 \ll 0$ and $f(0) - 0 \gg 0$. Applying the theorem to the function $f(x) - x$, we find x_0 such that $f(x_0) - x_0 \approx 0$. □

Note that the theorem and the corollary also follow if f only satisfies (1.22) for x and y of weight at most some infinite ω_1.

3.1.2. *Riemann integration.* The next step in the construction of elementary calculus is the Riemann integral. In Darboux's approach, a function is integrable if the infimum of the upper sums equals the supremum of the lower sums. Although several supremum principles are provable in ERNA and its extensions (see [**28**] and theorem 1.100), they are not very suited for a Darboux-like integral, because the supremum of nonstandard objects like lower sums does not have sufficiently strong properties. Therefore, we adopt Riemann's original approach, defining the integral as the limit of Riemann sums over ever finer partitions.

1.82. Definition. A partition π of $[a, b]$ is a vector $(x_1, \ldots, x_{n+1}, t_1, \ldots t_n)$ such that $x_i \leq t_i \leq x_{i+1}$ for all $1 \leq i \leq n$ and $a = x_1$ and $b = x_{n+1}$. The number $\delta_\pi = \max_{1 \leq i \leq n}(x_{i+1} - x_i)$ is called the 'mesh' of the partition π. We call a partition 'infinitely fine' if its mesh is infinitesimal.

Riemann integration implies quantifying over all partitions of an interval, which, as such, is not a first order-operation. However, encoding hyperfinite sets to hypernatural numbers, we are left with quantifying over all hypernaturals. The pairing function defined in section 2 is not suited for that purpose, because its iterations grow too fast for ERNA. Instead, we will use the pairing function

$$\pi^{(2)}(x, y) = \frac{1}{2}(x + y)(x + y + 1) + y.$$

All its iterations

$$\pi^{(n)}(x_1, \ldots x_n) := \pi(\pi^{(n-1)}(x_1, \ldots, x_{n-1}), x_n).$$

are available in ERNA, as one readily verifies by induction that

$$\|\pi^{(n)}(x_1, \ldots, x_n)\| \leq 2^{2^{n+1}} \|x_1, \ldots, x_n\|^{2^n},$$

for all x_i and hypernatural $n > 2$. Similarly, the decoding function $\left(\pi^{(n)}\right)^{-1}$, which yields the vector $(x_1, \ldots, x_n)$ when applied to $\pi^{(n)}(x_1, \ldots, x_n)$, can be defined in ERNA. Thus, ERNA allows quantification over all partitions of an interval.

1.83. Definition (Riemann Integration). Let f be a function defined on $[a, b]$.

(1) The Riemann sum corresponding to a partition $(x_1, \ldots, x_{n+1}, t_1, \ldots, t_n)$ is defined as $\sum_{i=1}^{n} f(t_i)(x_{i+1} - x_i)$.
(2) The function f is called 'Riemann integrable on $[a, b]$' if all Riemann sums of infinitely fine partitions are finite and infinitely close to each other. If so, the Riemann sum corresponding to the infinitely fine partition π of $[a, b]$ is denoted by $\int_a^b f(x)\, d_\pi x$.

1.84. Theorem. *A function, continuous and finite over* $[a, b]$*, is Riemann integrable over that interval.*

Proof. Let f be as stated and consider two infinitely fine partitions π_1 and π_2 of $[a, b]$. Let $\sum_{i=1}^{\omega_1} f(t_i)(x_{i+1} - x_i)$ and $\sum_{i=1}^{\omega_2} f(s_i)(y_{i+1} - y_i)$ be the respective Riemann sums. Using ERNA's definition by cases, we modify π_1 in the following way: if $[x_i, x_{i+1}]$ contains some y_j, break it into subintervals $[x_i, y_j]$ and $[y_j, x_{i+1}]$ and rename these subintervals to $[z_i, z_{i+1}]$ and $[z_{i+1}, z_{i+2}]$. Thus, the entry $f(t_i)(x_{i+1} - x_i)$ of the Riemann sum of π_1 is replaced by $f(t'_{i+1})(z_{i+2} - z_{i+1}) + f(t'_i)(z_{i+1} - z_i)$ with $t'_{i+1} := t_i$ and $t'_i := t_i$. Proceeding in the same way for π_2, we convert the original Riemann sums into $\sum_{i=1}^{\omega_3} f(t'_i)(z_{i+1} - z_i)$ and $\sum_{i=1}^{\omega_3} f(s'_i)(z_{i+1} - z_i)$, which share the upper bound ω_3 and the partition points. As all indices i and j are bounded by

$\omega_1+\omega_2+2$, this procedure is compatible with ERNA's definition by cases. Also, by construction, $t'_i \approx s'_i$. Hence, we have

$$\begin{aligned}\sum_{i=1}^{\omega_1} f(t_i)(x_{i+1}-x_i) - \sum_{i=1}^{\omega_2} f(s_i)(y_{i+1}-y_i) \\ &= \sum_{i=1}^{\omega_3} f(t'_i)(z_{i+1}-z_i) - \sum_{i=1}^{\omega_3} f(s'_i)(z_{i+1}-z_i) \\ &= \sum_{i=1}^{\omega_3} (f(t'_i)-f(s'_i))(z_{i+1}-z_i). \end{aligned} \tag{1.25}$$

Let $\overline{\varepsilon}$ be the maximum of the $|f(t'_{i-1})-f(s'_{i-1})|$ for $2 \leq i \leq \omega_3$, as provided by ERNA's explicit max operator. Because f is continuous over $[a,b]$, we have $\overline{\varepsilon} \approx 0$ and so the absolute value of (1.25) is at most $\sum_{n=1}^{\omega_3} \overline{\varepsilon}(z_i - z_{i-1}) = \overline{\varepsilon}(b-a) \approx 0$. Thus, the Riemann sums considered are infinitely close to each other. By theorem 1.77, the function f is finitely bounded on $[a,b]$ and hence every Riemann sum is in absolute value at most the finite number $(M+1)(b-a)$, with M as in (1.23). $\square$

3.1.3. *Differentiation.* Another key element of analysis is the derivative, defined in this paragraph. For brevity, we write '$\Delta_h f(x)$' for $\frac{f(x+h)-f(x)}{h}$.

1.85. Definition. [Differentiability] A function f is 'differentiable over (a,b)' if $\Delta_\varepsilon f(x) \approx \Delta_{\varepsilon'} f(x)$ is finite for all nonzero $\varepsilon, \varepsilon' \approx 0$ and all $a \ll x \ll b$.

If f is differentiable over (a,b) and $\varepsilon \approx 0$ is nonzero, then $\Delta_\varepsilon f(x_0)$ is called the 'ε-derivative of f at x_0' and denoted by $f'_\varepsilon(x_0)$. Any $f'_\varepsilon(x_0)$ with nonzero $\varepsilon \approx 0$ is a representative of 'the' derivative $f'(x_0)$, which is only defined up to infinitesimals. Thus, any statement about $f'(x_0)$ should be interpreted as a statement about $\Delta_\varepsilon f(x_0)$, quantified over all nonzero $\varepsilon \approx 0$.

A weaker notion than differentiability is provided by

1.86. Definition. [S-differentiability] A function f is called 'S-differentiable over (a,b)' if $\Delta_\varepsilon f(x) \approx \Delta_{\varepsilon'} f(x)$ is finite for all large enough $\varepsilon, \varepsilon' \approx 0$ and all $a \ll x \ll b$.

The informal expression 'for all large enough infinitesimals' in the previous definition is short for the external Σ_2-statement

$$\begin{aligned}&(\exists \varepsilon_0 \approx 0)(\forall \varepsilon, \varepsilon' \approx 0)(\forall x) \\ &\qquad \big[a \ll x \ll b \wedge |\varepsilon_0| < |\varepsilon|, |\varepsilon'| \to \Delta_\varepsilon f(x_0) \approx \Delta_{\varepsilon'} f(x_0)\big].\end{aligned} \tag{1.26}$$

The derivative is defined in the same way as for definition 1.85. A crucial point is that ε_0 does not depend on the choice of x. Indeed, otherwise ε_0 would be a function of x, i.e. in (1.26) the quantifier '$(\forall x)$' would be at the front. However, in ERNA, it is not possible to compute the function $\varepsilon_0(x)$ from the latter formula, as it involves '$\approx$'. In this case, the derivative would not be defined properly as it cannot be computed in a straightforward way.

Furthermore, 'S-differentiable' is short for 'standardly differentiable', and it does imply the classical ε-δ-definition of *uniform* differentiability, as we show in the next theorem. Thus, as in the case of continuity, the uniform version of differentiability is more natural than the pointwise one. Such phenomenon also occurs in the setting of Constructive Mathematics and in section 4, we discuss a possible connection to the latter. A more utilitarian argument in favour of S-differentiability is that it arises naturally in the proof of ERNA's version of the first fundamental theorem of calculus and Peano's existence theorem.

1.87. Theorem. *For f, S-differentiable over (c,d), we have, for $c \ll a \ll b \ll d$,*

$$(\forall^{st} k)(\exists^{st} N)(\forall^{st} h, h')(\forall^{st} x \in [a,b]) \\ \big[0 < |h|, |h'| < 1/N \to |\Delta_h f(x) - \Delta_{h'} f(x)| < 1/k\big]. \tag{1.27}$$

Proof. Choose ε_0 as in (1.26) and fix an infinite hypernatural ω_1. Then

$$(\forall h, h')(\forall x \in [a,b])\big[|\varepsilon_0| < |h|, |h'| \leq 1/N \\ \wedge \|h, h', x\| \leq \omega_1 \to \Delta_h f(x) \approx \Delta_{h'} f(x)\big]$$

holds for all infinite hypernatural N. Fixing $k \in \mathbb{N}$, we have in particular

$$(\forall h, h')(\forall x \in [a,b])\big[|\varepsilon_0| < |h|, |h'| \leq 1/N \\ \wedge \|h, h', x\| \leq \omega_1 \to |\Delta_h f(x) - \Delta_{h'} f(x)| < 1/k\big],$$

for all infinite hypernatural N. By corollary 1.47, the previous formula is equivalent to a quantifier-free one. Underflow yields

$$(\forall h, h')(\forall x \in [a,b])\big[|\varepsilon_0| < |h|, |h'| \leq 1/N \\ \wedge \|h, h', x\| \leq \omega_1 \to |\Delta_h f(x) - \Delta_{h'} f(x)| < 1/k\big],$$

for all $N \geq N(k) \in \mathbb{N}$, implying (1.27). □

Since (1.27) is stronger than pointwise differentiability, our derivative will have stronger properties, as witnessed by the following theorem. A function is said to be 'continuous over (a,b)' if it satisfies (1.22) for all $a \ll x, y, \ll b$.

1.88. Theorem. *If f is differentiable over (a,b), then $f'(x)$ is cont. over (a,b).*

Proof. Choose points $x \approx y$ such that $a \ll x < y \ll b$. If $|x-y| = \varepsilon \approx 0$, then

$$\Delta_\varepsilon f(x) = \tfrac{f(x+\varepsilon)-f(x)}{\varepsilon} = \tfrac{f(y)-f(y-\varepsilon)}{\varepsilon} = \tfrac{f(y-\varepsilon)-f(y)}{-\varepsilon} = \Delta_{-\varepsilon} f(y) \approx \Delta_\varepsilon f(y),$$

and thus $f'_{\varepsilon'}(x) \approx f'_\varepsilon(x) \approx f'_\varepsilon(y) \approx f'_{\varepsilon'}(y)$, for all nonzero $\varepsilon' \approx 0$. □

The theorem generalizes to S-differentiable functions, but not in an elegant way.

1.89. Corollary. *If f is S-differentiable over (a,b), then $f'_\varepsilon(x)$ is continuous over (a,b), for $\varepsilon \approx 0$ large enough.*

Proof. Let $\varepsilon_0 > 0$ be as in (1.26). Choose $x \approx y$ such that $a \ll x < y \ll b$. First, suppose $|x-y| = \varepsilon \geq \varepsilon_0$. The same proof as in the theorem yields this case. Now suppose $|x-y| = \varepsilon < \varepsilon_0$ and define $z = y + 2\varepsilon_0$. Then $|z-x| = \varepsilon' \geq \varepsilon_0$ and $|z-y| = \varepsilon'' \geq \varepsilon_0$ and by the previous case, we have $f'_{\varepsilon'}(x) \approx f'_{\varepsilon'}(z)$ and $f'_{\varepsilon''}(z) \approx f'_{\varepsilon''}(y)$. By the definition of S-differentiability, we have $f'_{\varepsilon'}(z) \approx f'_{\varepsilon''}(z)$, and thus $f'_{\varepsilon'''}(x) \approx f'_{\varepsilon'}(x) \approx f'_{\varepsilon''}(y) \approx f'_{\varepsilon'''}(y)$, for all $\varepsilon''' \geq \varepsilon_0$. □

Since the derivative is only defined up to infinitesimals in ERNA, the statement $f'(x) > 0$ is not very strong, as $f'(x) \approx 0$ may also hold. Similarly, $f(x) < f(y)$ is consistent with $f(x) \approx f(y)$ and we need stronger forms of inequality to express meaningful properties of functions and their derivatives.

1.90. Definition. A function f is $\ll$-increasing over an interval $[a,b]$, if for all $x, y \in [a,b]$ we have $x \ll y \to f(x) \ll f(y)$. Likewise for $\ll$-decreasing.

1.91. Theorem. *If f is differentiable over (a,b), there is an $N \in \mathbb{N}$ such that*

(1) *if $f'(x_0) \gg 0$, then f is $\ll$-increasing in $[x_0 - \frac{1}{N}, x_0 + \frac{1}{N}]$,*
(2) *if $f'(x_0) \ll 0$, then f is $\ll$-decreasing in $[x_0 - \frac{1}{N}, x_0 + \frac{1}{N}]$,*

for all $a \ll x_0 \ll b$.

Proof. For the first item, $f'(x_0) \gg 0$ implies $f(y) > f(z)$ for all y, z satisfying $y, z \approx x_0$ and $y > z$. Fix an infinite number ω_1 and let $M \gg 0$ be $f'(x_0)/2$. By the previous, the following sentence is true for all infinite hypernaturals N

$$(\forall y, z)\big[\|y,z\| \le \omega_1 \wedge y > z \wedge |x_0 - z| < \tfrac{1}{N} \wedge |x_0 - y| < \tfrac{1}{N} \to f(y) > f(z) + M(y-z)\big].$$

By corollary 1.47, the previous formula is equivalent to a quantifier-free one. Applying underflow yields the first item, as f is continuous over (a,b). Likewise for the second item. □

We have previously pointed out that ERNA proves many theorems of basic analysis with equality replaced by $\approx$. However, the formula $x \approx y$ is equivalent to $(\forall^{st} k)(|x - y| < \frac{1}{k})$, i.e. '$x \approx y$' is not a Δ_0-formula. Similarly, in constructive analysis, equality is a (strict) Π_1-statement. Thus, we can expect there to be a connection between constructive analysis (see also section 4.2) and our results. In this way, the constructive notion of inequality '$<$' then corresponds to '$\ll$'.

A function is said to be 'continuous at a' if (1.22) holds for $x = a$.

1.92. Theorem (Mean value theorem). *If f is differentiable over (a,b) and continuous in a and b, then there is an $x_0 \in [a,b]$ such that $f'(x_0) \approx \frac{f(b)-f(a)}{b-a}$.*

Proof. Let f be as in the theorem. First, we prove the particular case where $f(a) \approx f(b)$. By theorem 1.77, f attains its maximum (up to infinitesimals), say in x_0, and its minimum (idem), say in x_1, over $[a,b]$. If $f(x_0) \approx f(x_1) \approx f(a)$, then f is constant up to infinitesimals. By theorem 1.91 we have $f'(x) \approx 0$ for all $a \ll x \ll b$. If $f(x_0) \not\approx f(a)$, then by theorem 1.91 we have $f'(x_0) \approx 0$. The case $f(x_1) \not\approx f(a)$ is treated in a similar way. The general case can be reduced to the particular case by using the function $F(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$. □

3.1.4. *The first fundamental theorem of calculus.* In this paragraph, we obtain ERNA's version of the first fundamental theorem of calculus.

1.93. Definition. If π is an infinitely fine partition of $[a,b]$, we denote by $\underline{x}$ the least partition point not exceeding x. If f is integrable over $[a,b]$, we define

$$F_\pi(x) := \int_a^{\underline{x}} f(t)\, d_\pi t. \tag{1.28}$$

In ERNA, there is no standard-part function mapping a finite number x to the unique standard number $r \approx x$. Consequently, there is no natural way to avoid that integrals are only defined up to infinitesimals. The same occurred in ERNA's predecessor NQA$^+$ proposed in [**11**]. There, differentiation and integration cancel each other out on the condition that the mesh du of the hyperfinite partition and the infinitesimal y used in the derivative are related by $du/y \approx 0$. This requirement is hidden under a complicated definition of the integral (see [**11**, Axiom 18]). Our definitions of integration and differentiation are quite natural and we still obtain an elegant version of the first Fundamental Theorem of Calculus, see corollary 1.95 below.

1.94. THEOREM. *Let f be continuous on $[a,b]$. For every $\eta \approx 0$ and every hyperfine partition π of $[a,b]$ with $\delta_\pi/\eta \approx 0$, we have $\Delta_\eta F_\pi(x) \approx f(x)$ for all $a \ll x \ll b$.*

PROOF. Let f, π and η be as stated and fix $a \ll x_0 \ll b$. Then we have

$$F_\pi(x_0) = \sum_{i=1}^{\omega_1} f(t_i)(x_{i+1} - x_i) \text{ and } F_\pi(x_0 + \eta) = \sum_{n=1}^{\omega_2} f(t_i)(x_{i+1} - x_i),$$

with $\omega_2 > \omega_1$. Now, let M be the largest and m the smallest of the $f(t_i)$ for $\omega_1 < i \leq \omega_2$. Then $F_\pi(x_0 + \eta) - F_\pi(x_0)$ equals

$$\sum_{i=\omega_1+1}^{\omega_2} f(t_i)(x_{i+1} - x_i) \leq M \sum_{i=\omega_1+1}^{\omega_2} (x_{i+1} - x_i) = M(x_{\omega_2+1} - x_{\omega_1+1}).$$

By definition 1.93, $|x_0 - x_{\omega_1+1}| \leq \delta_\pi$ and $|(x_0 + \eta) - x_{\omega_2+1}| \leq \delta_\pi$. Consequently

$$\eta - 2\delta_\pi < x_{\omega_2+1} - x_{\omega_1+1} < \eta + 2\delta_\pi,$$

which implies that $\frac{x_{\omega_2+1} - x_{\omega_1+1}}{\eta} \approx 1$. Therefore, $\frac{1}{\eta}(F_\pi(x_0 + \eta) - F_\pi(x_0)) \lessapprox M$. Combining with the similar result for m, we obtain

$$m \lessapprox \frac{F_\pi(x_0 + \eta) - F_\pi(x_0)}{\eta} \lessapprox M.$$

Assuming that $M = f(t_{j_1})$ and $m = f(t_{j_2})$, we have $m \approx f(x_0) \approx M$ thanks to continuity and $t_{j_1} \approx t_{j_2} \approx x_0$. Hence, $\Delta_\eta F_\pi(x_0) \approx f(x_0)$. □

The previous theorem can be formulated much more elegantly if we see $\int_a^b f(x)\,dx$ and $F'(x)$ as ERNA objects 'defined up to infinitesimals' (compare [**1**, §5]). Accordingly, we interpret an informal statement about $\int_a^b f(x)\,dx$ as a statement about all the Riemann sums corresponding to infinitely fine partitions of $[a,b]$. As we can quantify over all partions of an interval, this informal statement can be expressed in the language of ERNA and we will sometimes forget the distinction between informal and formal terminology. With these conventions, the previous theorem becomes.

1.95. COROLLARY (First Fundamental Theorem of Calculus). *Let f be a continuous function on $[a,b]$ and assume $F(x) = \int_a^x f(t)dt$. Then F is S-differentiable on $[a,b]$ and $F'(x) \approx f(x)$ holds for all $a \ll x \ll b$*

PROOF. Observe that the choice of η in the theorem does not depend on x. □

On a philosophical note, we mention that it seems impossible to develop basic analysis in ERNA (or in any classical nonstandard system without a standard-part function) in a quantifier-free way. Indeed, to study the function $F(x)$, we cannot use the quantifier-free definition of differentiability, but we have to revert back to the (standard) non-quantifier-free definition. The same holds for Peano's existence theorem in ERNA, but equally for nonstandard set theory, e.g. the treatment of the nonstandard representative $\frac{\varepsilon}{\varepsilon^2+x^2}$ ($\varepsilon \approx 0$) of the Dirac delta function. In chapter II, we suggest possible solutions for this problem.

Although the proof of theorem 1.95 may seem straightforward, the condition $\delta_\pi/\eta \approx 0$ is highly non-constructive (see also (1.26)) and cannot be 'read off' from the first fundamental theorem of calculus. Thus, it seems only fair to say that this theorem, at the very least, does not agree with the spirit of finitism. However, the conditions of the first fundamental theorem can be weakened to remove this problem. Indeed, if $\delta \approx 0$, then η such that $\eta^2 \geq \delta$ is easily computed and satisfies $\delta/\eta \approx 0$. However, there are many more of these conditions and none of them is optimal.

3.1.5. *Differential equations.* In this paragraph we prove ERNA's version of the Peano existence theorem for ordinary differential equations. In [**50**], Sommer and Suppes sketch a proof of this theorem inside ERNA. Their sketch is based on the classical stepwise construction of the function $\phi(x)$ which, in the limit, satisfies the necessary properties. This construction is a prime example of the elegance of Nonstandard Analysis (see [**26**]) and we will carry out this construction explicitly inside ERNA in the proof of the following theorem.

1.96. Theorem (Peano existence theorem). *Let $f(x,y)$ be continuous on the rectangle $|x| \leq a$, $|y| \leq b$, let M be a finite upper bound for $|f|$ there and let $\alpha = \min(a, b/M)$. Then there is a function ϕ, S-differentiable for $|x| \leq \alpha$, such that*

$$\phi(0) = 0 \text{ and } \phi'(x) \approx f(x, \phi(x)). \tag{1.29}$$

Proof. Without loss of generality, we may assume $a = b = \alpha = 1$. We will only consider positive x, the proof for negative x is analogous.

First, define $x_k := k/\omega = k\varepsilon$ for $k \leq \omega$ and

$$y_m := \sum_{k=1}^{m} f(x_{k-1}, y_{k-1})\varepsilon \text{ and } \phi(x) := \sum_{m=1}^{\omega} T_\psi(m,x) y_m, \tag{1.30}$$

where $\psi(m,x) \equiv (x_{m-1} < x \leq x_m)$ and $y_0 = 0$. It is an easy verification that the function $\phi(x)$ is available in ERNA. We verify that $\phi(x)$ satisfies the conditions of the theorem. Fix $0 \ll x \ll 1$ and a nonzero positive infinitesimal η such that $\varepsilon/\eta \approx 0$ for $\varepsilon = \frac{1}{\omega}$. The case for negative η is treated similarly. Now assume that

$$x_{\omega_1 - 1} < x + \eta \leq x_{\omega_1} \text{ and } x_{\omega_2 - 1} < x \leq x_{\omega_2} \tag{1.31}$$

for certain numbers $\omega_2 \leq \omega_1 \leq \omega$. Then $\phi(x+\eta) = y_{\omega_1}$ and $\phi(x) = y_{\omega_2}$ and

$$\begin{aligned}\phi(x+\eta) - \phi(x) = y_{\omega_1} - y_{\omega_2} &= \sum_{k=1}^{\omega_1} f(x_{k-1}, y_{k-1})\varepsilon - \sum_{k=1}^{\omega_2} f(x_{k-1}, y_{k-1})\varepsilon \\ &= \varepsilon \sum_{k=\omega_2+1}^{\omega_1} f(x_{k-1}, y_{k-1})\end{aligned}$$

Assume $f(x_N, y_N)$ ($f(x_M, y_M)$, respectively) is the largest (the least, respectively) of all $f(x_i, y_i)$ for i between ω_1 and $\omega_2 + 1$. Define $M' = \omega_1 - \omega_2 - 1$; there holds

$$\varepsilon M' f(x_M, y_M) \leq \phi(x+\eta) - \phi(x) \leq \varepsilon M' f(x_N, y_N)$$

and also

$$\frac{\varepsilon}{\eta} M' f(x_M, y_M) \leq \Delta_\eta \phi(x) \leq \frac{\varepsilon}{\eta} M' f(x_N, y_N). \tag{1.32}$$

By the definition of x_n, there holds

$$x_{\omega_1 - 1} - x_{\omega_2} = \tfrac{\omega_1 - 1}{\omega} - \tfrac{\omega_2}{\omega} = \tfrac{\omega_1 - \omega_2 - 1}{\omega} = \varepsilon M'. \tag{1.33}$$

But (1.31) implies $x_{\omega_1-1} - x_{\omega_2} < \eta$, which yields $\frac{\varepsilon}{\eta}M' < 1$. Again by the definition of x_n, there holds

$$x_{\omega_1} - x_{\omega_2-1} = \tfrac{\omega_1}{\omega} - \tfrac{\omega_2-1}{\omega} = \tfrac{\omega_1-\omega_2+1}{\omega} = \varepsilon M' + 2\varepsilon. \tag{1.34}$$

But (1.31) also implies $x_{\omega_1} - x_{\omega_2-1} > \eta$, which yields $\frac{\varepsilon}{\eta}M' > 1 - 2\varepsilon/\eta$. Together with $\frac{\varepsilon}{\eta}M' < 1$, proved above, this yields $\frac{\varepsilon}{\eta}M' \approx 1$. It is clear that $x \approx x_N \approx x_M$. We now prove that $y_M \approx y_N \approx \phi(x)$. Then (1.32) and the continuity of f imply

$$\Delta_\eta \phi(x) \approx f(x_N, y_N) \approx f(x_M, y_M) \approx f(x, \phi(x)) \tag{1.35}$$

and we are done.

Assume that $N < M$; the case $N > M$ is treated analogously. From (1.30), there follows

$$y_N - y_M = \sum_{k=1}^{N} f(x_{k-1}, y_{k-1})\varepsilon - \sum_{k=1}^{M} f(x_{k-1}, y_{k-1})\varepsilon = \varepsilon \sum_{k=N+1}^{M} f(x_{k-1}, y_{k-1}).$$

By the Weierstraß extremum theorem, f is bounded on $[0,1]$, say by $M'' \in \mathbb{N}$. Then (1.30) implies

$$y_N - y_M \leq \varepsilon(M - N - 1)M'' \leq \varepsilon(\omega_1 - \omega_2 - 1)M'' = \varepsilon M' M''.$$

By the previous, this implies $y_M \approx y_N$. In the same way, $y_N \approx y_{\omega_2} = \phi(x)$. $\square$

In [**50**], Sommer and Suppes claim that $\phi(x)$, as defined in (1.30), is differentiable (in the sense of definition 1.85). However, due to the absence of a 'standard-part function', the function $\phi(x)$ defined in (1.30) remains piecewise constant, be it on the infinitesimal level. Thus, if η is too small, we have $\phi(x_0) = \phi(x_0+\eta)$ for some x_0 and hence $\Delta_\eta \phi(x_0) = 0$, even if $\phi(x)$ is strictly increasing. Hence, it is obvious that $\phi(x)$ cannot be differentiable, but only S-differentiable. Thus, the Peano existence theorem implicitly involves a condition $\varepsilon/\eta \approx 0$ similar to the condition $\delta_\pi/\eta \approx 0$ in the fundamental theorem of calculus. As in the latter, S-differentiability hides this technical requirement, but this does not change the fact that ε-δ-like formulas occur.

Before we continue, we point out that the theorems proved so far fall in either of two fundamentally different classes. A good representative of the first kind is Weierstraß's extremum theorem: it fails when we limit the weight of x and y to ω in (1.22). Also, the consequent of this theorem is a property of *all* numbers in $[a, b]$ of arbitrary depth. On the other hand, the Brouwer fixed point theorem does go through with the aforementioned limitation and its consequent only asserts the existence of a number x_0 of a certain depth. However, if we were to require a fixed point of arbitrary depth, the resulting 'uniform' fixed point theorem becomes part of the first class. The distinction made here will turn out to be essential in the section 'ERNA and Reverse Mathematics'.

3.2. Mathematics with Transfer. In this section, we prove some well-known results from ordinary mathematics in ERNA + Π_1-TRANS. By theorem 1.73, we are allowed to use bar transfer.

3.2.1. *Completeness.* In this paragraph, we prove an ERNA-version of Cauchy and Dedekind completeness, to be understood 'up to infinitesimals'. Indeed, Cauchy completeness is well-known to be equivalent to ACA_0 over RCA_0 and the theory ACA_0 has the same first-order strength as PA (see [**46**] for details).

We first treat Cauchy completeness. An everywhere defined function $\tau(n)$, not involving min, is called a 'sequence'.

1.97. Definition. A sequence $\tau(n)$ is 'Cauchy' if

$$(\forall^{st} k)(\exists^{st} N)(\forall^{st} n, m)\left(n, m \geq N \to |\tau(n) - \tau(m)| < \tfrac{1}{k+1}\right). \tag{1.36}$$

If a is a constant, we say that a sequence $\tau(n)$ is 'convergent to a' if

$$(\forall^{st} k)(\exists^{st} N)(\forall^{st} n)\left(n \geq N \to |\tau(n) - a| < \tfrac{1}{k+1}\right).$$

Clearly, the constant a is only unique up to infinitesimals. We have the following theorem, provable in $\mathrm{ERNA} + \Pi_1$-TRANS.

1.98. Theorem (Cauchy completeness). *Let $\tau(n)$ be a near-standard Cauchy sequence. Then all terms of infinite index are infinitely close to each other and $\tau(n)$ is standard convergent to any of these.*

Proof. If $\tau(n)$ is as required, (1.36) holds. In this formula, fix any $k \in \mathbb{N}$ and find $N \in \mathbb{N}$ such that

$$(\forall^{st} n, m)\left(n, m \geq N \to |\tau(n) - \tau(m)| < \tfrac{1}{k+1}\right). \tag{1.37}$$

In $\mathrm{ERNA} + \Pi_1$-TRANS, this implies

$$(\forall n, m)\left(n, m \geq N \to |\tau(n) - \tau(m)| \lessapprox \tfrac{1}{k+1}\right), \tag{1.38}$$

which shows that $\tau(n)$ is convergent to $\tau(m)$ for any infinite m. Since (1.38) can be derived for all $k \in \mathbb{N}$, we have $|\tau(n) - \tau(m)| \lessapprox \frac{1}{k+1}$ for all infinite hypernaturals n, m and $k \in \mathbb{N}$. Hence, infinitely indexed terms differ by infinitesimals. □

Note that since (1.37) involves parameters k and N, we cannot use Π_1-TRANS$^-$ here. However, by theorem 1.130, the latter principle is not useless. Below, we need the following version of the previous theorem, provable in ERNA.

1.99. Corollary. *Let $\tau(n)$ be an internal Cauchy sequence. There is an infinite hypernatural m_0 such that all $\tau(m)$ are infinitely close to each other for all infinite $m \leq m_0$, and $\tau(n)$ is convergent to any of these.*

Proof. Use overflow to obtain (1.38) with the quantifier '$(\forall n, m)$' bounded by the infinite number $\overline{m}(k)$. By theorem 1.55, the latter is infinite for all k up to some infinite number ω_2. Let ω_3 be the least $\overline{m}(k)$ for $k \leq \omega_2$ (see theorem 1.53). Thus, (1.38) holds for $n, m \leq \omega_3$ and we have obtained the theorem for $m_0 = \omega_3$. □

Next, we treat Dedekind completeness. In particular, we prove the following supremum principle in $\mathrm{ERNA} + \Pi_1$-TRANS. A preliminary version of it restricted to particular formulas is to be found in [**27**].

1.100. Theorem (Supremum Principle). *Let b be a finite constant and $\varphi(x)$ a quantifier-free formula of L^{st}, such that*

(i) *$\varphi(x)$ holds for no $x > b$*

(ii) $\varphi(x)$ *holds for at least one finite* x.

Then there is a constant β*, given by an explicit* ERNA*-formula, not involving* min*, with the following properties:*

(iii) $\varphi(x)$ *holds for no* $x \gg \beta$

(iv) *for every finite* $\varepsilon \gg 0$ *there are rational* $x > \beta - \varepsilon$ *such that* $\varphi(x)$ *holds.*

The several constants β *satisfying these requirements differ by infinitesimals.*

Proof. By definition, the number β must be in $[a, b]$ and we can approximate β by inductively dividing the interval $[a, b]$ in two, testing one for the presence of β and proceeding with the subinterval with contains β. It is a long and technical verification that this can be done in ERNA + Π_1-TRANS. See [**28**, Theorem 67] for full details. □

Note that the theorem is limited to standard formulas, as the formulation for near-standard formulas is too involved. In ERNA plus Π_2-TRANS (Π_3-TRANS), we can prove a version of the theorem where φ is in Π_1 (Π_2). However, the proof becomes unmanageable.

3.2.2. *Continuity.* In this paragraph, we introduce the well-known ε-δ definition of (uniform) continuity in ERNA. This will have immediate consequences for the continuity, differentiability and integration results obtained earlier.

1.101. Definition. *A function* $f(x)$ *is called 'S-continuous over* $[a, b]$*' if*

$$(\forall^{st} k)(\exists^{st} N)(\forall^{st} x, y \in [a, b])(|x - y| < 1/N \to |f(x) - f(y)| < 1/k) \tag{1.39}$$

The following theorem shows that continuity implies S-continuity for internal functions.

1.102. Theorem. *In* ERNA*, continuity, i.e.* (1.22)*, implies S-continuity, i.e.* (1.39)*, for any internal* $f(x)$.

Proof. Assume that (1.22) holds for an internal function $f(x)$. Fix $k \in \mathbb{N}$ and consider the following internal formula $\Phi(n)$

$$(\forall x, y)\big((\|x, y\| \leq \omega \wedge |x - y| < 1/n) \to |f(x) - f(y)| < 1/k\big).$$

By theorem 1.47, the formula $\Phi(n)$ is equivalent to a quantifier-free one. By assumption, $\Phi(n)$ holds for all infinite n. Hence, by underflow, there is an $N \in \mathbb{N}$ such that $(\forall n \geq N)\Phi(n)$. From this, (1.39) follows immediately. □

The following theorem shows that S-continuity implies continuity for near-standard functions, if Π_1-transfer is available.

1.103. Theorem. *In* ERNA + Π_1-TRANS, (1.39) *implies* (1.22) *for near-standard functions.*

Proof. Let $f(x)$ be near-standard and S-continuous over $[a, b]$. Fix nonzero $k \in \mathbb{N}$ and let $N \in \mathbb{N}$ be such that

$$(\forall^{st} x, y \in [a, b])(|x - y| < 1/N \to |f(x) - f(y)| < 1/k). \tag{1.40}$$

By bar transfer, we obtain

$$(\forall x, y \in [a, b])(|x - y| < 1/N \to |f(x) - f(y)| \lessapprox 1/k).$$

In particular, $|f(x) - f(y)| \lessapprox 1/k$ if $x \approx y$ for $x, y \in [a, b]$. But $k \in \mathbb{N}$ can be chosen arbitrarily large and hence $f(x) \approx f(y)$ if $x \approx y$ for $x \in [a, b]$. □

The previous theorem has the interesting consequence that all the theorems we obtained in the section 'Mathematics without Transfer' now follow for 'continuous' replaced with 'S-continuous' and 'internal' replaced with 'near-standard'. Thus, we know that Π_1-transfer is sufficient to prove these theorems. In section 4, we show that this transfer principle is exactly what is needed to prove many of these theorems, i.e. Π_1-transfer is also necessary.

For completeness, we note that in ERNA the formula (1.40) implies (1.22) for x and y of weight at most some infinite ω_1. This is easily proved via overflow in the same way as in corollary 1.99. Thus, ERNA proves the Intermediate value theorem and the Brouwer fixed point theorem (see theorem 1.80 and corollary 1.81).

3.2.3. *Separation.* Here, we prove ERNA's version of Σ_1-Separation (see [**46**, I.11.7 and IV.4.4]). Although ERNA's language does not contain set variables, we can simulate subsets of $\mathbb{N}$ in the following way. Let $(x)_y$ be the function which calculates the power of the $(y+1)$-th prime number in the prime decomposition of x. It is an easy verification that this function is available in ERNA. Thus, we write '$m \in M$' if $(M)_m > 0$ and in this way, subsets of $\mathbb{N}$ can be mimicked in ERNA (compare [**32**]). The proof takes place in ERNA + Π_1-TRANS.

1.104. Theorem (Σ_1^{st}-Separation). *For $i = 1, 2$, let ψ_i be formulas $(\exists^{st} m)\varphi_i(m, n)$ with $\varphi_i \in L^{st}$ quantifier-free. If $(\forall^{st} n)(\neg\psi_1(n) \vee \neg\psi_2(n))$, then*

$$(\exists M)(\forall^{st} n)\left[\psi_1(n) \to n \in M \wedge \psi_2(n) \to n \notin M\right].$$

Proof. Let φ_i and ψ_i be as stated. Define $T(n)$ as true if $(\exists m \leq \omega)\varphi_1(m, n)$ and false otherwise. By theorem 1.47, the formula $(\exists m \leq \omega)\varphi(m, n)$ is equivalent to a quantifier-free one. By theorem 1.45, the internal function $T(n)$ is well-defined. By Σ_1-transfer, $(\exists m \leq \omega)\varphi(m, n)$ is equivalent to $(\exists^{st} m)\varphi(m, n)$, if n is finite. Thus, for finite n, $T(n) = 1$ if and only if $(\exists^{st} m)\varphi(m, n)$.

It is an easy verification that the function 'p_k = the k-th prime number' is available in ERNA. Now define the number $M := \prod_{n=1}^{\omega} p_n^{T(n-1)}$. By Π_1-transfer, it is clear that M has the right properties. □

3.2.4. *The Isomorphism Theorem.* In this paragraph, we prove an upgraded version of the Isomorphism Theorem (see [**49**, Section 6]) in ERNA + Π_1-TRANS. This theorem states that for a finite set of internal atomic propositions in ERNA's language, we can replace each occurrence of 'ω', 'ε' and '$x \approx y$' with, respectively, 'n_0', '$1/n_0$' and '$|x - y| < 1/b$' (for some $n_0, b \in \mathbb{N}$) in such a way that the propositions remain true. We first prove the Isomorphism Theorem and then discuss its philosophical implications. We also discuss why the following definition is natural in this context.

1.105. Definition. An ERNA-term $\tau(\vec{x})$ is called 'intensional' if there is a $k \in \mathbb{N}$ such that $(\forall \vec{x})\big[\|\tau(\vec{x})\| > \log^k\left(\|\vec{x}\|\right)\big]$.

The function $\log^k n$ is defined as $(\mu m \leq n)(2_k^m > n)$. Intensional objects are also discussed in section 5.3.

1.106. Theorem (Isomorphism Theorem). *Let $\mathcal{T}$ be a finite set of* intensional constant terms *of* ERNA*, not including* min *and closed under subterms. There is an isomorphism f from $\mathcal{T}$ to a finite set of rationals such that*

(i) *$f(0) = 0$, $f(1) = 1$ and $f(\omega) = n_0$, for some $n_0 \in \mathbb{N}$,*

(ii) $f(g(\tau_1,\dots,\tau_k)) = g(f(\tau_1),\dots,f(\tau_k))$, *for all non-atomic terms in* $\mathcal{T}$,
(iii) $\tau \approx 0$ *iff* $|f(\tau)| < \frac{1}{b}$, *for some* $n_0 > b \in \mathbb{N}$,
(iv) τ *is infinite iff* $|f(\tau)| > b$, *for some* $n_0 > b \in \mathbb{N}$,
(v) τ *is hypernatural iff* $f(\tau)$ *is natural,*
(vi) $\sigma \le \tau$ *iff* $f(\sigma) \le f(\tau)$.

Proof. Let $\mathcal{T}$ be as in the theorem and let D be the maximum depth of the terms in $\mathcal{T}$. Complete $\mathcal{T}$ with terms $\lfloor\tau\rfloor$ for $\tau \in \mathcal{T}$, if necessary. By theorem 1.33, there is a $B_1 \in \mathbb{N}$ such that $\|h(\vec{x})\| \le 2_{B_1}^{\|\vec{x}\|}$ for all terms h in $\mathcal{T}$. As all terms in $\mathcal{T}$ are assumed intensional, there is a $B_2 \in \mathbb{N}$ such that $\|h(\vec{x})\| > \log^{B_2}(\|\vec{x}\|)$ for all terms h in $\mathcal{T}$. Let B be the maximum of B_1 and B_2 and add the term $\log^B \omega$ to $\mathcal{T}$ if necessary.

Then, define Ψ as the conjunction of all true formulas $\mathcal{N}(\tau)$, $\sigma = \tau$ and $\sigma \le \theta$ with $\tau, \sigma, \theta \in \mathcal{T}$. Let $\Psi(m)$ be Ψ with all occurrences of ω replaced with the free variable m. As $\varepsilon = \frac{1}{\omega}$, any occurrence of ε in Ψ is replaced with $\frac{1}{m}$. By construction, there holds $\Psi(\omega)$. As ω is infinite and 2^1_{2DB} is finite, this implies $(\exists m > 2^1_{2DB})\Psi(m)$. By Σ_1-transfer, there holds $(\exists^{st} m > 2^1_{2DB})\Psi(m)$, i.e. there is a *finite* number m such that $m > 2^1_{2DB}$ and $\Psi(m)$. Let m_0 be such a number. Then, let f be any map which maps ω to m_0 and has property (ii). By construction, f satisfies (v) and (vi). To conclude, we show that f also satisfies (iii) and (iv). First of all, by theorem 1.33, if $\tau \in \mathcal{T}$ does not involve ω, then it satisfies $\|\tau\| \le 2^1_{BD}$ and hence τ must be finite. Thus, if $\tau \in \mathcal{T}$ is infinite, it must involve ω. Hence, we have $\tau = \sigma(\omega)$ for some term $\sigma \in \mathcal{T}$ and as all terms in $\mathcal{T}$ are intensional, we have $|\lfloor\sigma(n)\rfloor| \ge \log^B n$. In particular, we have, for $\tau > 0$,

$$f(\lfloor\tau\rfloor) = f(\lfloor\sigma(\omega)\rfloor) = \lfloor\sigma(f(\omega))\rfloor = \lfloor\sigma(m_0)\rfloor \ge \log^B m_0$$

Thus, if $\tau > 0$ is infinite, then $f(\lfloor\tau\rfloor) \ge \log^B m_0$, which implies $f(\tau) \ge \log^B m_0$. Hence, for all infinite $\tau \in \mathcal{T}$, we have $|f(\tau)| \ge \log^B m_0$. Now assume that $|f(\tau)| \ge \log^B m_0$ for some $\tau \in \mathcal{T}$. This yields $|f(\tau)| \ge f(\log^B \omega)$ and, by item (vi), there holds $|\tau| \ge \log^B \omega$. Thus, τ is infinite and we have proved item (iv) for $b = \log^B m_0$. As item (iv) implies item (iii), we are done. □

In comparison to Sommer and Suppes' approach, we removed the 'reasonably sound' condition from the Isomorphism Theorem, which is a significant improvement (compare [**49**, Theorem 6.1]), and we fixed its proof. However, we added the 'intensionality' condition and it may not be clear why this condition is natural. We give several arguments, both heuristic and formal.

First of all, the best-known example of a *non-intensional* function is $\log^* n = (\mu k \le n)(\log^k n \le 1)$. It can be computed that for $n_0 = 2^{65536}$, which is larger than the number of particles in the universe, $\log^* n_0$ is at most five. Thus, for practical purposes, $\log^* n$ may be regarded as an eventually constant function. Moreover, in Chapter II we prove theorem 2.75 (see section 5.3) which states that there are models of ERNA in which $\log^* n$ *is* an eventually constant function. Since the Isomorphism Theorem is intended to deal with models of physical problems, it seems reasonable to choose a model of ERNA which corresponds to the real world, i.e. one where $\log^* n$ is eventually constant. Alternatively, one can exclude $\log^* n$ from the Isomorphism theorem, replacing it with a constant if necessary. Secondly, another interpretation of theorem 2.75 shows that ERNA cannot prove anything about non-intensional terms. Thus, we might as well exclude them from the Isomorphism Theorem, as we cannot learn anything about them in ERNA anyway. Nonetheless, the Isomorphism Theorem turns a negative result (theorem 2.75) into a positive one.

We now discuss the philosophical implications of the Isomorphism Theorem.

First of all, it shows that the use of irrational numbers (and functions taking such values) in Physics is merely a convenient calculus tool. Indeed, let $\mathcal{M}$ be a model of a (necessarily finite) physical problem $\mathcal{P}$ that involves irrational numbers. We can approximate these numbers by hyperrationals with infinitesimal precision. After replacing the irrational numbers with these approximations, we apply the Isomorphism Theorem to obtain a model $\mathcal{M}'$ of $\mathcal{P}$ that only involves *rational* numbers. We second Sommer and Suppes' claim that 'the continuum may be real for Platonists, but it can nowhere be unequivocally identified in the real world of physical experiments.' (see [**49**, Introduction]).

Secondly, the representation of physical quantities such as space and time as continuous variables is called into question by the Isomorphism Theorem. Indeed, by the latter, a discrete set of rational numbers already suffices to model a physical problem and hence no physical experiment can decide the 'true' nature (discrete or continuous) of physical quantities. The obvious 'human-all-too-human' way to avoid the previous 'undecidability' result, is to simply state that one does not accept the Isomorphism Theorem (or Π_1-transfer) and hence one is not bound to its implications. We counter with the following observation: in section 4 we show that Π_1-transfer is equivalent to the 'continuity principle' which states that ε-δ continuity implies nonstandard continuity. The latter formalizes the heuristic notion of continuity, which is fundamental in the informal reasoning inherent to applied sciences, especially Physics. Thus, the continuity principle is inherent to Physics and so is Π_1-transfer and the Isomorphism Theorem.

Thus, we obtain our boutade: *Whether reality is continuous or discrete is undecidable because of the way mathematics is used in Physics.*

4. Reverse Mathematics in ERNA

4.1. A copy of Reverse Mathematics for $\mathbf{WKL}_0$. In this section, we prove the equivalences between Π_1-transfer and the theorems of ordinary mathematics listed in theorem 1.3. Most of the latter are derived from theorems equivalent to Weak König's lemma (see theorem 1.2 and [**46**]) by replacing equality with '$\approx$'. Hence, the Reverse Mathematics of ERNA + Π_1-TRANS is a 'copy up to infinitesimals' of the Reverse Mathematics for WKL_0.

We also mention Strict Reverse Mathematics (SRM), recently introduced by Harvey Friedman, which is 'a form of Reverse Mathematics relying on no coding mechanisms, where every statement considered must be strictly mathematical'. Comparing the usual definition of continuity with [**46**, Definition II.6.1], it is clear that Reverse Mathematics uses significant coding machinery. In contrast, ERNA can approximate most functions that appear in mathematical practice by near-standard functions and bar transfer enables us to prove many well-known results and the associated reversal, all with minimal coding. Thus, the Reverse Mathematics of ERNA + Π_1-TRANS is also a contribution to SRM.

Recall that we allow standard parameters in Π_1-TRANS (see after schema 1.57). In the same way, we always allow standard parameters in the principles enumerated in theorem 1.3. Inspecting e.g. formula (1.107), it is clear why we have to allow parameters in those principles. See also theorem 1.130.

4.1.1. *Completeness.* Recall theorem 1.98 which expresses that ERNA's field is (Cauchy) complete 'up to infinitesimals'. Thus, in the context of ERNA, we refer to this theorem as the 'Cauchy completeness principle'. We have the following theorem.

1.107. THEOREM. *In* ERNA*,* Π_1*-transfer is equivalent to the Cauchy completeness principle.*

PROOF. By theorems 1.73 and 1.98, the forward implication is immediate. To obtain the reverse implication, assume the Cauchy completeness principle, let φ be as in Π_1-TRANS and assume $\varphi(m)$ for $m \in \mathbb{N}$. Let $\tau(n)$ be a near-standard Cauchy sequence and define

$$\sigma(n) = \begin{cases} \tau(n) & (\forall m \leq n)\varphi(m) \\ n & \text{otherwise} \end{cases}. \tag{1.41}$$

By definition 1.66, $\sigma(n)$ is also near-standard. By assumption $\sigma(n) = \tau(n)$ for $n \in \mathbb{N}$ and hence $\sigma(n)$ is also a Cauchy sequence. By Cauchy completeness, we have $\sigma(k) \approx \sigma(k')$ for all infinite k, k'. If $\sigma(k) = k$ for some infinite k, then also $\sigma(k+1) = k+1$, by (1.41). But then $\sigma(k) \not\approx \sigma(k+1)$, which yields a contradiction. Thus, for all infinite k, there must hold $\sigma(k) = \tau(k)$. By (1.41), this implies $\varphi(m)$ for all m and hence Π_1-TRANS follows. □

Note that without theorem 1.73, the Cauchy completeness principle would be limited to standard sequences, which excludes e.g. hyperrational approximations of sequences of reals.

For those interested in minimal proofs, we mention that, by theorem 1.45, the statement that 'If a binary sequence $\tau(n) \in L^{st}$ is zero for $n \in \mathbb{N}$, it is zero everywhere.' is equivalent to Π_1-transfer. Thus, the proof of theorem 1.107 can be reduced to a short, but meaningless, proof. Since we believe that proofs are more than meaningless 'games' with symbols, we do not explore this further. Furthermore, many results in this thesis are difficult, if not impossible, to discover using such 'minimalist' techniques. Also, the aforementioned statement does not occur in mathematical practice. Finally, as there is no function $T_{\overline{\varphi}}(n)$ for near-standard quantifier-free formulas φ (see theorem 1.45), it is not clear how to obtain bar transfer.

4.1.2. *Continuity.* Consider the following 'continuity principle' (see theorem 1.103).

1.108. PRINCIPLE (Continuity principle). *For near-standard functions, the definition of S-continuity implies that of continuity, i.e.* (1.39) *implies* (1.22)*.*

1.109. THEOREM. *In* ERNA*, the continuity principle is equivalent to* Π_1-TRANS*.*

PROOF. The reverse implication is immediate from theorem 1.103. Conversely, assume the continuity principle and consider a quantifier-free formula φ of L^{st}, such that $\varphi(n)$ holds for $n \in \mathbb{N}$. Let f be near-standard and S-continuous over $[a, b]$. By cases, we define the near-standard function

$$g(x) = \begin{cases} f(x) & (\forall n \leq \|x\|)\varphi(n) \\ \|x\| & \text{otherwise} \end{cases}. \tag{1.42}$$

For standard x, we have $\|x\| \in \mathbb{N}$ and $(\forall n \leq \|x\|)\varphi(n)$ holds by assumption. Hence, for standard x, $g(x) = f(x)$, the latter being a function S-continuous over $[a, b]$. Thus, $g(x)$ is S-continuous over $[a, b]$ too and, by assumption, this implies that $g(x)$ is continuous over $[a, b]$. Now suppose there is an infinite k such that $\neg\varphi(k)$ and let k_0 be the least number with this property. Fix $a \ll x_0 \ll b$ with weight $\leq k_0$. Assume $k_1 \geq k_0$ is prime. By (A.143), $\|x_0 + 1/k_1\| \geq k_1$ and thus we have $g(x_0 + 1/k_1) = \|x + 1/k_1\|$, where the latter is infinite. But by assumption $g(x)$ is continuous, which implies $g(x_0) \approx g(x_0 + 1/k_1)$, as $x_0 \approx x_0 + 1/k_1$. Since $g(x_0) = f(x_0)$, the latter is a finite number by corollary 1.79. This yields a contradiction and hence $\varphi(n)$ holds for all n. This implies Π_1-TRANS □

Note that the theorem still holds if we only require f to be continuous over (a, b) in the continuity principle. We will refer to this as the 'continuity principle' too.

In the previous proof, ERNA's weight function $\|x\|$ is used not as a proof theoretic tool (as in the consistency proof of [**49**]), but as an ERNA-function that is everywhere discontinuous. However, from the proof of the theorem, it is clear that we could replace $\|x\|$ by a function which has a jump in its graph for some $a \ll x_0 \ll b$. Indeed, in the proof, we only consider continuity for $a \ll x_0 \ll b$. Also, this alternative function g obviates any claims that theorem 1.3 is not meaningful from the point of view of mathematical practice because the function defined in (1.42) would somehow be artificial.

Now consider the following version of Weierstraß' extremum theorem.

1.110. Principle (Weierstraß extremum principle). *If f is near-standard and S-continuous over $[a, b]$, there is a number $c \in [a, b]$ such that for all $x \in [a, b]$, we have $|f(x)| \lessapprox |f(c)|$.*

1.111. Theorem. *In* ERNA*, the Weierstraß extremum principle is equivalent to* Π_1-TRANS*.*

Proof. The reverse implication is immediate from theorems 1.77 and 1.103. Conversely, assume the Weierstraß extremum principle and consider a quantifier-free formula φ of L^{st} such that $\varphi(n)$ is valid for all $n \in \mathbb{N}$. Define $g(x)$ as in (1.42). In the same way as in the previous proof, g is S-continuous over $[a, b]$ and by the Weierstraß extremum principle there is a number $c \in [a, b]$ such that $|g(x)| \lessapprox |g(c)|$, for all $x \in [a, b]$. Now suppose there is an n_0 such that $\neg\varphi(n_0)$. By theorem 1.51, there is an $a \ll x_0 \ll b$ with weight at least $1 + \lceil \max\{n_0, |g(c)|\} \rceil$. As $\|x_0\| > n_0$, this implies $(\exists n \leq \|x_0\|)\neg\varphi(n)$ and by the definition of g, we have $|g(x_0)| = \|x_0\|$. But by the definition of x_0, we have $|g(x_0)| = \|x_0\| \gg |g(c)|$. This is a contradiction and hence $\varphi(n)$ must hold for all n, which implies Π_1-TRANS. □

Note that the proof can be easily adapted to a weaker version of Weierstraß' extremum theorem where $|f(x)|$ is only bounded by some $M \in \mathbb{N}$ for $x \in [a, b]$.

Next, we treat Brouwer's fixed point theorem. We need the following definition.

1.112. Definition. The point x_0 is a 'fixed point up to infinitesimals' of f if $f(x_0) \approx x_0$.

After theorem 1.103, we noted that in ERNA every S-continuous $[0,1] \to [0,1]$-function has a fixed point up to infinitesimals. As RCA_0 proves the one-dimensional Brouwer fixed point theorem, this supports our claim concerning the ressemblance between the Reverse Mathematics of WKL_0 and that of ERNA$+\Pi_1$-TRANS. However, the following strengthening of Brouwer's fixed point theorem is not provable in ERNA.

1.113. Principle (Uniform Brouwer fixed point principle). *For every* $[0,1] \to [0,1]$-*function* f*, near-standard and S-continuous over* $[0,1]$*, there is a fixed point up to infinitesimals of arbitrary weight.*

1.114. Theorem. *In* ERNA*, the Uniform Brouwer fixed point principle is equivalent to* Π_1-TRANS*.*

Proof. The reverse implication, is immediate from theorem 1.103 and the Brouwer fixed point theorem (see corollary 1.81). Conversely, assume the Uniform Brouwer fixed point principle and consider a quantifier-free formula φ of L^{st} such that $\varphi(n)$ is true for all $n \in \mathbb{N}$. Define $g(x)$ as in (1.42). In the same way as in the previous proofs, g is S-continuous over $[0,1]$. Now suppose $\neg\varphi(n_0)$ for some infinite n_0. By the Uniform Brouwer fixed point principle, there is a point $x_0 \in [0,1]$ with weight at least n_0 such that $g(x_0) \approx x_0$. If $g(x_0)$ equals $\|x_0\|$, we have $\|x_0\| \approx x_0$, which is obviously false. Thus, we have $g(x_0) = f(x_0)$, which implies $(\forall n \leq \|x_0\|)\varphi(n)$, by definition. As $\|x_0\| \geq n_0$, this yields $\varphi(n_0)$, which contradicts $\neg\varphi(n_0)$. Thus, $\varphi(n)$ holds for all n and we obtain Π_1-TRANS. □

4.1.3. *Integration and differentiability.* First, we consider the following principle concerned with Riemann integration.

1.115. Principle (Riemann integration principle). *A near-standard function which is S-continuous over* $[a,b]$*, is Riemann integrable there.*

1.116. Theorem. *In the theory* ERNA*, the Riemann integration principle is equivalent to* Π_1-TRANS*.*

Proof. The reverse implication is immediate from theorems 1.84 and 1.103. Conversely, assume that the Riemann integration principle holds and consider a quantifier-free formula φ of L^{st} such that $\varphi(n)$ is true for all $n \in \mathbb{N}$. Let $g(x)$ be as in (1.42). As $\varphi(n)$ is true for all $n \in \mathbb{N}$, we have $g(x) = f(x)$ for all standard x and hence the Riemann integration principle applies to g. Overflow applied to $(\forall^{st} n)\varphi(n)$ yields $(\forall n \leq \omega_1)\varphi(n)$ for some infinite ω_1. Hence, $g(x) = f(x)$ for all x such that $\|x\| \leq \omega_1$. Then put $\omega_2 = \lfloor \omega_1/2 \rfloor - 2$ and consider the equidistant partition with mesh $1/\omega_2$ and points $t_i = \frac{x_{i+1}+x_i}{2}$. As $\|t_i\| \leq \omega_1$, it is clear that $g(t_i) = f(t_i)$ for $1 \leq i \leq \omega_2$ and assume the Riemann sum of f corresponding to this partition is the finite number S.

Now suppose there is a (necessarily infinite) hypernatural n_1 such that $\neg\varphi(n_1)$ and let $n_0 \geq n_1$ be prime. By the definition of $g(x)$, there follows $g(x) = \|x\|$ if $\|x\| \geq n_0$. Then consider the equidistant partition with mesh $1/n_0$ and points $t_i = \frac{x_{i+1}+x_i}{2}$. The corresponding Riemann sum is easily calculated:

$$\sum_{i=1}^{n_0} g(t_i)(x_i - x_{i-1}) = \sum_{i=1}^{n_0} \|t_i\| \; \frac{1}{n_0} \geq \sum_{i=1}^{n_0} n_0 \; \frac{1}{n_0} = n_0.$$

By the Riemann integration principle, there holds $S \approx n_0$. Obviously, this is impossible and the assumption that there is a number n_1 such that $\neg\varphi(n_1)$ is false. This implies Π_1-TRANS and we are done. $\square$

Theorem 1.109 suggests an alternative proof for the reverse implication. Indeed, assume the Riemann integration principle and suppose there are $x_0, y_0 \in [a,b]$ such that $x_0 \approx y_0$ and $f(x_0) \not\approx f(y_0)$. Assume $x_0 < y_0$ and fix an infinitely fine partition π of $[a,b]$ for which $x_i < x_0 < y_0 < x_{i+1}$ and $x_0 = t_i$ for some i. Change π into π' by putting $y_0 = t_i$. Then the corresponding Riemann sums differ a noninfinitesimal amount and we have have a contradiction. Thus, f is continuous and theorem 1.109 implies Π_1-transfer.

It should be noted that the format of the continuity principle, i.e. 'standard definition' implies 'nonstandard definition', in many cases results in a principle equivalent to Π_1-TRANS. Indeed, the statement 'If a (near)-standard function satisfies (1.27) then it is differentiable on (a,b)' is also equivalent to Π_1-TRANS. Similar statements can be found based on the definition of near-standard term (1.10) and (1.17), the definition of Riemann integration or even the notion of a modulus of 'uniform differentiability'. However, these principles do not really qualify as a part of mathematical practice or ordinary mathematics, in contrast to the continuity principle.

Next, we consider the following version of the first fundamental theorem of calculus.

1.117. Principle (FTC$_1$). *Let f be near-standard and S-continuous on $[a,b]$ and assume $F(x) = \int_a^x f(t)dt$. Then F is S-differentiable on $[a,b]$ and $F'(x) \approx f(x)$ holds for all $a \ll x \ll b$.*

We have the following theorem.

1.118. Theorem. *In* ERNA*,* FTC$_1$ *is equivalent to* Π_1-TRANS*.*

Proof. The reverse implication is immediate by corollary 1.95 and theorem 1.103. For the forward implication, assume FTC$_1$ and let f be as stated there. By FTC$_1$, $F(x)$ is S-differentiable over (a,b) and hence $F'(x)$ is continuous over (a,b), by corollary 1.89. Again, by FTC$_1$, the formula $F'(x) \approx f(x)$ holds for all $a \ll x \ll b$ and hence $f(x)$ is also continuous over (a,b). By theorem 1.109, this implies Π_1-TRANS and we are done. $\square$

Consider the following version of the Peano existence theorem.

1.119. Principle (Peano existence principle). *Let $f(x,y)$ be near-standard and S-continuous on the rectangle $|x| \le a$, $|y| \le b$, let M be a finite upper bound for f there and let $\alpha = \min(a, b/M)$. Then there is a function ϕ, S-differentiable for $|x| < \alpha$, such that*

$$\phi(0) = 0 \text{ and } \phi'(x) \approx f(x, \phi(x)).$$

1.120. Theorem. *In* ERNA*, the Peano existence principle is equivalent to* Π_1-TRANS*.*

Proof. The reverse implication is immediate by theorem 1.96 and theorem 1.103. For the forward implication, we prove that the function $\phi'(x)$ is continuous in the same way as for FTC$_1$. Thus, $f(x, \phi(x))$ is continuous over (a,b). From this, Π_1-TRANS follows in the same way as for theorem 1.109. $\square$

4.1.4. *Approximation and Bernstein polynomials.* In this paragraph, we study ERNA's version of the Weierstraß approximation theorem. The latter is equivalent to WKL (see [**46**, Theorem IV.2.5]).

1.121. Definition. For a function f, define the n-th Bernstein polynomial as

$$B_n(f)(x) := \sum_{k=0}^{n} f(k/n)\binom{n}{k}x^k(1-x)^{n-k}.$$

1.122. Principle (Weierstraß approximation principle). *Let f be near-standard and S-cont. on $[a,b]$. Then $B_m(f)(x) \approx f(x)$ for all $x \in [a,b]$ and infinite m.*

1.123. Theorem. *In* ERNA*, the Weierstraß approximation theorem is equivalent to* Π_1-TRANS.

Proof. Assume Π_1-TRANS. In [**14**], an elementary, rather tedious, proof of the Weierstraß approximation theorem is given, based on Bernstein's original proof. This proof can easily be adapted to the context of ERNA to prove

$$(\forall^{st} k)(\exists^{st} N)(\forall^{st} n \geq N)(\forall^{st} x \in [a,b])\big(|B_n(f)(x) - f(x)| < 1/k\big).$$

Applying bar transfer to the innermost universal formula implies $B_m(f)(x) \approx f(x)$ for all $x \in [a,b]$ and all infinite m.

Now assume the Weierstraß approximation theorem and let f be as stated there. Its a technical verification that ERNA proves that $B_m(f)(x)$ is continuous on $[a,b]$ for small enough infinite m. Since $B_m(f)(x) \approx f(x)$ for all $x \in [a,b]$ and infinite m, this implies the continuity of f on $[a,b]$ and theorem 1.109 yields Π_1-TRANS. $\square$

4.1.5. *Modulus of uniform continuity.* In this paragraph, we study ERNA's version of the 'modulus of uniform continuity' (see [**46**, Definition IV.2.1]). The statement 'every uniform continuous function has a modulus of uniform continuity' is equivalent to WKL over RCA_0 ([**46**, IV.2.9]).

1.124. Definition. Let f be a function defined on $[a,b]$. A function $h(k,m)$ is a modulus of uniform continuity for f if for all m we have

$$\begin{aligned}(\forall^{st} k)(\forall x, y \in [a,b])\big[&\|x,y\| \leq m \\ &\wedge |x-y| < \tfrac{1}{h(k,m)} \to |f(x) - f(y)| < \tfrac{1}{k+1}\big],\end{aligned} \tag{1.43}$$

and $h(k,m)$ is finite for finite k.

Note that this definition is weaker than the usual one. Indeed, our modulus depends on $\|x,y\|$. Alternatively, one can say that there is a modulus for every initial segment of the hypernaturals. These insights turn out to be crucial for ERNA's version of the Bolzano-Weierstraß theorem, proved in section 4.3.

1.125. Principle (Modulus principle). *Every near-standard function, S-conti-nuous on $[a,b]$, has a modulus of uniform continuity.*

1.126. Theorem. *In* ERNA*, the modulus principle is equivalent to* Π_1-TRANS.

Proof. First, assume Π_1-TRANS and let f be as in the modulus principle. Then,

$$(\forall^{st} k)(\exists^{st} N)(\forall^{st} x, y \in [a,b])\big[|x-y| < \tfrac{1}{N} \to |f(x) - f(y)| < \tfrac{1}{k+2}\big],$$

and by bar transfer

$$(\forall^{st}k)(\exists^{st}N)(\forall x, y \in [a,b])\big[|x-y| < \tfrac{1}{N} \to |f(x)-f(y)| \lessapprox \tfrac{1}{k+2}\big],$$

and also

$$(\forall^{st}k)(\exists^{st}N)(\forall x, y \in [a,b])\big[|x-y| < \tfrac{1}{N} \to |f(x)-f(y)| < \tfrac{1}{k+1}\big].$$

Thus, for any fixed m, there holds

$$(\forall^{st}k)(\exists^{st}N)(\forall x, y \in [a,b])\big[\|x,y\| \leq m \wedge |x-y| < \tfrac{1}{N} \to |f(x)-f(y)| < \tfrac{1}{k+1}\big].$$

By corollary 1.47, the innermost universal formula may be treated as quantifier-free. Then define $h(k,m)$ as

$$(\mu N \leq \omega)(\forall x, y \in [a,b])\big[\|x,y\| \leq m \wedge |x-y| < \tfrac{1}{N} \to |f(x)-f(y)| < \tfrac{1}{k+1}\big],$$

which is a suitable modulus.

For the forward implication, assume the modulus principle and let f be as stated there. Then f satisfies (1.43) for some modulus $h(k,m)$. Now fix $x_0, y_0 \in [a,b]$ such that $x_0 \approx y_0$ and apply (1.43) for $m_0 = \|x_0, y_0\|$. This implies that $f(x_0) \approx f(y_0)$ and hence f is also continuous over $[a,b]$. By theorem 1.109, Π_1-TRANS follows and we are done. □

Similarly, we could define a modulus of convergence or a modulus of equicontinuity (see [**46**, p. 110]) and the associated versions of the modulus principle would also be equivalent to Π_1-TRANS. Also, the modulus principle is a special case of the following general principle.

1.127. Principle (Π_3-modulus). *Let φ be standard and quantifier-free. If*

$$(\forall^{st}k)(\exists^{st}N)(\forall^{st}n)\varphi(k,N,n),$$

then there is a function $\alpha(k,M)$ such that for all M

$$(\forall^{st}k)(\forall n \leq M)\varphi(k,\alpha(k,M),n).$$

Obviously, this principle is equivalent to Π_1-TRANS but special cases such as the modulus principle are more interesting.

4.1.6. *Conclusion.* We have concluded the proof of theorem 1.3 and we repeat our dictum.

The Reverse Mathematics of ERNA + Π_1-TRANS *is a 'copy up to infinitesimals' of the Reverse Mathematics of* WKL_0.

Moreover, our results have several philosophical implications for mathematics and Physics, specifically regarding physical and mathematical modelling.

4.1.7. *Future research.* We list our most ambitious goals first. We believe that there exists a version of theorem 1.3 for Π_2 and Π_3-transfer. The predicate '$\approx$' would be replaced by arbitrarily good approximation, i.e. we would have an *arbitrarily small* infinitesimal error. We also suspect that many theorems of constructive/computable mathematics/[**46**] can be directly translated to ERNA. We have many examples, but require some more time to develop this further.

Obviously, the list in theorem 1.3 is not exhaustive and many more theorems equivalent to WKL_0 are expected to have a version which is equivalent to Π_1-TRANS. We list two examples of such theorems.

First, we point to [**56**] where Keita Yokoyama proves the equivalence between WKL and Cauchy's integral theorem which states that a complex function $f \in C^1(\Omega)$

satisfies the well-known zero-law $\oint_\gamma f(z)\,dz = 0$ for a sufficiently well-behaved closed curve $\gamma \subset \Omega$. It is beyond the scope of this thesis to develop complex analysis in ERNA, but we mention that Π_1-TRANS is equivalent to an ERNA-version of the Cauchy integral theorem with 'approximate' zero-law $\oint_\gamma f(z)\,dz \approx 0$.

Similarly, the Jordan curve theorem is equivalent to WKL_0 ([**58**]) and ERNA's version of the former theorem only implies that for every arc $A(x)$ with endpoints in the interior and exterior of the Jordan curve $J(x)$, there is a point x_0 such that $A(x_0) \approx J(x_0)$. Thus, the Jordan curve and the arc only meet 'up to infinitesimals', consistent with our dictum. Also, the locus of $J(x)$ is its infinitesimal neighbourhood and the condition that the interior of $J(x)$ is bounded, gives rise to Π_1-transfer, in the same way as for the Weierstraß extremum principle. To prove the Jordan curve theorem in ERNA+Π_1-TRANS, construct a polygon P with $\omega-1$ vertices $J(i/\omega)$. Then P and J are infinitely close everywhere and the Jordan curve theorem for polygons is straightforward.

Also, it is not inconceivable that a natural version of Σ_1^{st}-separation (see theorem 1.104) is equivalent to Π_1-TRANS.

Finally, it should be noted that one slight anomaly is present in theorem 1.3: the Cauchy completeness property is equivalent to ACA over RCA_0 ([**46**, Theorem III.2.2]), but ERNA's version of Cauchy completeness is equivalent to Π_1-TRANS, which is not in accordance with our dictum. In section 4.2, we give a possible explanation for this phenomenon, which may be a fruitful avenue of research. Thus, rather than sweeping the anomaly that Cauchy completeness presents under the proverbial carpet, we embrace it, in accordance with the tenets of Good Science as promoted by Richard Feynmann (see e.g. [**16**]).

4.1.8. *Finitistic Reverse Mathematics.* In this paragraph, we obtain a finitistically acceptable version of theorem 1.3. Indeed, by theorem 1.75, Π_1-TRANS is too strong for finitistic mathematics and hence all principles enumerated in theorem 1.3 are too. By theorem 1.58, Π_1-TRANS$^-$, the parameter-free version of Π_1-TRANS, *is* suitable for finitistic mathematics. However, in the context of Cauchy sequences and continuity, we have always applied transfer to formulas *with* parameters (see e.g. the proofs of theorems 1.98 and 1.103). Thus, Π_1-TRANS$^-$ is not a suitable replacement for Π_1-TRANS. This ceases to be true if we use a slightly stronger version of continuity, defined next.

1.128. Definition. *A function $f(x)$ is called 'M-continuous over $[a,b]$' if there is a standard function h such that*

$$(\forall^{st} k)(\forall^{st} x, y \in [a,b])(|x-y| < \tfrac{1}{h(k)} \to |f(x)-f(y)| < \tfrac{1}{k}). \tag{1.44}$$

Note that (1.44) is the 'skolemized' version of (1.39). Many functions that appear in mathematical practice are M-continuous. M-continuity also plays an important role in constructive analysis.

1.129. Principle (MC). *For standard functions, the definition of M-continuity implies that of continuity, i.e.* (1.44) *implies* (1.22).

We do not allow standard parameters in the functions of MC. We have the following theorem.

1.130. Theorem. *In* ERNA, MC *is equivalent to* Π_1-TRANS$^-$.

Proof. For the inverse implication, note that (1.44) is universal and parameter-free. Applying Π_1-TRANS$^-$, we obtain

$$(\forall k)(\forall x, y \in [a,b])(|x-y| < \tfrac{1}{h(k)} \to |f(x)-f(y)| < \tfrac{1}{k}).$$

As $h(k)$ is standard, it is finite for finite k. Thus, we see that $x \approx y$ implies $|f(x)-f(y)| < \frac{1}{k}$ for all $x, y \in [a,b]$ and finite k. This immediately implies (1.22).

For the forward implication, proceed as in the proof of theorem 1.109, except that f and g are M-continuous in (1.42). □

Presumably, a parameter-free version of bar transfer can be derived from the schema Π_1-TRANS$^-$. This would generalize MC to near-standard functions and we could weaken the notion of M-continuity by allowing near-standard functions h instead of only standard ones.

Similarly, we could formulate a version of the Cauchy property involving a modulus function h. The convergence of such Cauchy sequences would be equivalent to Π_1-TRANS$^-$.

As most equivalences in theorem 1.3 are proved using the continuity principle (see 1.108), it is almost immediate that replacing S-continuity with M-continuity in theorem 1.3, yields a list of theorems equivalent to Π_1-TRANS$^-$. The latter list qualifies for 'finitistic' Reverse Mathematics.

Given Hilbert's stance on intuitionism, it is somewhat ironic that M-continuity, a concept from constructive analysis, provides the key to finitistic Reverse Mathematics.

4.2. ERNA and Constructive Reverse Mathematics. In this section, we speculate on the connection between the Reverse Mathematics for ERNA + Π_1-TRANS and Constructive Reverse Mathematics. We first briefly introduce the latter.

Constructive mathematics ([**5**–**7**]) is described by Douglas Bridges as 'that mathematics which is characterized by *numerical content* and *computational method.*' ([**6**, p. 1]). Thus, in constructive mathematics, the quantifier '$(\exists x)$' means 'there is an algorithm to compute the object x'. This is stronger than the 'ideal' notion of existence in the sense of Plato used in classical mathematics. From the constructive perspective, the law of excluded is suspect since it carries non-constructive content and therefore it is excluded from constructive mathematics. Constructive Reverse Mathematics studies equivalences between both constructive and non-constructive theorems in a constructive base theory (see e.g. [**30**, **31**]). In the case of non-constructive theorems, one of the goals is to find out just how much of the law of excluded middle (or another non-constructive principle) is needed to prove such a theorem. In this context, the following principle occurs in relation to Cauchy completeness.

1.131. Principle (Σ_1-PEM). *For all quantifier-free φ, there holds*

$$(\exists n)\varphi(n) \vee (\forall n)\neg\varphi(n).$$

In the previous, the existential quantifier '$(\exists n)$' means that 'a number n can be computed'. Also, Π_1-transfer is equivalent to the following schema.

1.132. Principle (Σ_1-TRANS). *For all quantifier-free $\varphi \in L^{st}$, there holds*

$$(\exists^{st} n)\varphi(n) \vee (\forall n)\neg\varphi(n).$$

In this way, Σ_1-TRANS is a form of 'hyperexcluded middle': it excludes the possibility that $(\forall^{st} n)\varphi(n) \wedge (\exists n)\neg\varphi(n)$. Not only does Σ_1-transfer ressemble Σ_1-PEM, we can also easily compute a witness to $(\exists^{st} n)\varphi(n)$ by the number $(\mu n \leq \omega)\varphi(n)$. Thus, we see that Π_1-transfer has 'constructive' content, similar to that of Σ_1-PEM. As the latter is related to Cauchy completeness, it is no surprise that Π_1-transfer is also related to Cauchy completeness (see theorem 1.107).

We can take this analogy further by considering another principle from Constructive Reverse Mathematics related to Cauchy completeness.

1.133. Principle (Π^0_1-AC_{00}). *For $A \in \Pi_1$, we have*

$$(\forall m)(\exists n)A(m,n) \to (\exists \alpha)(\forall m)A(m,\alpha(m)).$$

In constructive mathematics, this choice principle implies that every Cauchy sequence has a modulus. In ERNA + Π_1-TRANS, we have the following theorem.

1.134. Theorem (Countable Universal Choice). *Let $A(m,n)$ be $(\forall^{st} k)B(k,m,n)$ with $B \in L^{st}$ and quantifier-free. Then $(\forall^{st} m)(\exists^{st} n)A(m,n)$ implies the formula $(\forall^{st} m)A(m,\alpha(m))$ for some nonstandard function α.*

Proof. Let $A(m,n)$ be as stated. By transfer, $(\forall^{st} m)(\exists^{st} n)A(m,n)$ implies the formula $(\forall^{st} m)(\exists^{st} n)(\forall k)B(k,m,n)$. This yields $(\forall^{st} m)(\exists^{st} n)(\forall k \leq \omega)B(k,m,n)$ and the function $\alpha(m) = (\mu n \leq \omega)(\forall k \leq \omega)B(k,m,n)$ is a suitable modulus. □

The previous theorem can be modified to be equivalent to Π_1-TRANS. Indeed, see principle 1.127.

Finally, we note that there is no internal bounded formula which is equivalent to $(\exists^{st} n)(\forall^{st} m)\varphi(n,m)$, even in the presence of Π_2-transfer. Thus, we cannot compute a witness to $(\exists^{st} n)(\forall^{st} m)\varphi(n,m)$ as we could for Σ_1-formulas. However, this problem disappears in the 'stratified' framework, see chapter III. By theorem 3.8 and the above, we expect the Reverse Mathematics for Σ_n-PEM to be similar to that for Σ_n-transfer.

4.3. Reverse Mathematics beyond WKL_0. In this section, we study theorems and theories related to Reverse Mathematics which take us beyond WKL_0 or the associated theory ERNA + Π_1-TRANS.

4.3.1. *The Bolzano-Weierstraß theorem.* In this paragraph, we study ERNA's version of the Bolzano-Weierstraß theorem and related theorems. In [**46**, III.2.2], the following theorem is listed.

1.135. Theorem. *The following assertions are pairwise equivalent over* RCA_0.

(1) ACA_0.
(2) The Bolzano/Weierstraß theorem*: Every bounded sequence of real numbers contains a convergent subsequence.*
(3) *Every Cauchy sequence of real numbers is convergent.*
(4) *Every bounded sequence of real numbers has a least upper bound.*
(5) The monotone convergence theorem*: Every bounded increasing sequence of real numbers is convergent.*

Thus, the Bolzano-Weierstraß theorem clearly goes beyond WKL_0. Below, in theorem 1.161, we obtain results similar to theorem 1.135. However, for the proofs in this paragraph, we repeatedly require specific instances of the external minimum schema of NQA^+. Rather than adding the entire external minimum schema to ERNA, we only add the following schema, called EXIT for 'external iteration'. A function is called *arithmetical* if it is weakly increasing in all its variables and does not involve min.

1.136. Axiom schema (EXIT). *For all arithmetical f, if $f(n,\omega)$ is finite for $n \in \mathbb{N}$, then $f^m(0,\omega)$ is finite for all $m \in \mathbb{N}$.*

Recall that '$f^n(x)$' denotes n applications of f to x, as defined in (1.3). Also, for notational convenience, we assume that the symbol 'ω' in $\tau(\vec{x},\omega)$ represents all occurrences of ω in $\tau(\vec{x},\omega)$, i.e. $\tau(\vec{x},m)$ is $\tau(\vec{x},\omega)$ with all occurrences of ω replaced with the new variable m.

We define ERNA^+ as ERNA plus the EXIT schema. The proof of the following theorem takes place in ERNA^+.

1.137. Theorem (Internal Subsequence principle). *For every internal $\tau(n)$, there is an explicit function $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\tau(\sigma(n))$ is monotone over $\mathbb{N}$.*

Proof. Assume $\tau(n)$ is as in the theorem and let $\psi(n)$ be the formula $(\forall^{st} m)(m > n \to \tau(m) \le \tau(n))$. The proof is divided in three parts.

First, assume $\neg\psi(n)$ holds for all $n \in \mathbb{N}$, i.e. we have

$$(\forall^{st} n)(\exists^{st} m)(m > n \wedge \tau(m) > \tau(n)). \tag{1.45}$$

Then define

$$f(k) = (\mu m \le \omega)(m > k \wedge \tau(m) > \tau(k)) \text{ and } \sigma(n) = f^n(1). \tag{1.46}$$

By example 1.37, the term $\sigma(n)$ is available in ERNA.

Second, assume there are only finitely many n such that $\psi(n)$ holds, i.e. there is a $k_0 \in \mathbb{N}$ such that for all $n \ge k_0$ we have $\neg\psi(n)$. Defining $f(k)$ as in (1.46) and $\sigma(n)$ as $f^n(k_0)$ concludes this case.

Third, assume there are infinitely many n such that $\psi(n)$. Hence, for all $k \in \mathbb{N}$ there is a natural number $n \ge k$ such that

$$(\forall^{st} m)(m > n \to \tau(m) \le \tau(n)). \tag{1.47}$$

Applying overflow yields a term $\overline{m}(k)$ which is infinite for $k \in \mathbb{N}$. By theorem 1.55, there is an infinite ω_2 such that $\overline{m}(k)$ is infinite for all $k \le \omega_2$. Let ω_3 be the least of all $\overline{m}(k)$ for $k \le \omega_2$. Then, we have

$$(\forall^{st} k)(\exists^{st} n \ge k)(\forall m \le \omega_3)(m > n \to \tau(m) \le \tau(n)).$$

and we define

$$f(k) = (\mu n \le \omega)\big[n \ge k \wedge (\forall m \le \omega_3)(m > n \to \tau(m) \le \tau(n))\big]. \tag{1.48}$$

The term $\sigma(n)$ is defined as in (1.46).

In each case, the fact that $\sigma(n)$ is finite for finite n is immediate from EXIT. $\square$

By overflow, we know that the term $\tau(\sigma(n))$ from the theorem has the same monotonous behaviour over a hyperfinite initial segment. However, for internal functions, this segment cannot be arbitrarily long. Indeed, let $\tau_1(n)$ and $\tau_2(n)$ be a strictly increasing and a strictly decreasing internal sequence and define

$$\tau(n) := \begin{cases} \tau_1(n) & n \leq \omega \\ \tau_2(n) & n > \omega \end{cases}. \tag{1.49}$$

Then $\tau(n)$ is an internal sequence and let σ be the function provided by the previous theorem. It is clear that $\tau(\sigma(n))$ can only be increasing for $n \leq \omega$.

In the presence of Π_2-transfer, we can obtain a stronger subsequence principle for (near-)standard sequences. Thus, the following theorem is proved in $\text{ERNA}^+ + \Pi_2$-TRANS.

1.138. Theorem (Standard Subsequence principle). *For every $\tau(n) \in L^{st}$, there is an explicit $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\tau(\sigma(n,\omega))$ is strictly increasing or weakly decreasing over $\mathbb{N}$. Also, for every M, there is an N such that $\tau(\sigma(n,N))$ is similarly monotone for $n \leq M$.*

Proof. Let τ be as stated, let σ be as in the proof of the previous theorem and assume we are in the first case. The other cases are treated analogously. Then the formula $(\forall^{st} n)(\tau(\sigma(n,\omega)) < \tau(\sigma(n+1,\omega)))$ implies $(\forall n \leq \overline{n}(\omega))(\tau(\sigma(n,\omega)) < \tau(\sigma(n+1,\omega)))$, by overflow. Thus, $\overline{n}(\omega)$ is infinite, which implies $(\forall^{st} k)(\overline{n}(\omega) > k)$ and also $(\forall^{st} k)(\exists m)(\overline{n}(m) > k)$, and finally $(\forall^{st} k)(\exists^{st} m)(\overline{n}(m) > k)$, by Σ_1-transfer. Tranfer for Π_2-formulas implies $(\forall k)(\exists m)(\overline{n}(m) > k)$ and as $\overline{n}(m)$ is the largest $n' \leq m$ such that $(\forall n \leq n')(\tau(\sigma(n,m)) < \tau(\sigma(n+1,m)))$, the theorem follows. □

Note that the proof of the theorem fails for *internal* sequences $\tau(n,\omega)$, because $\tau(n,N)$ is not always $\tau(n,N')$ for $N \neq N'$. However, by definition, near-standard functions $\tau(n,\omega)$ only vary infinitesimally when we change the parameter ω (see (1.10)) and hence we have the following definition and theorem. The latter is proved in $\text{ERNA}^+ + \Pi_2$-TRANS.

1.139. Definition. A sequence $\tau(n)$ is called '$\approx$-increasing' if $\tau(n) \lessapprox \tau(n+1)$, for all n. Similarly for '$\approx$-decreasing' and '$\approx$-monotone'.

1.140. Theorem (Near-Standard Subsequence principle). *For every near-standard $\tau(n,\omega)$, there is an explicit $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\tau(\sigma(n,\omega),\omega)$ is monotone over $\mathbb{N}$. Also, for every M, there is an N such that $\tau(\sigma(n,N),\omega)$ is similarly $\approx$-monotone for $n \leq M$.*

Proof. Let τ be as stated, let σ be as in the proof of theorem 1.137 and assume we are in the first case. The other cases are treated analogously. Then the formula

$$(\forall^{st} n)\big[\tau(\sigma(n,\omega),\omega) < \tau(\sigma(n+1,\omega),\omega)\big]$$

implies, by overflow

$$(\forall n \leq \overline{n}(\omega))\big[\tau(\sigma(n,\omega),\omega) < \tau(\sigma(n+1,\omega),\omega)\big].$$

Thus, $\overline{n}(\omega)$ is infinite, which implies $(\forall^{st} k)(\overline{n}(\omega) > k)$ and $(\forall^{st} k)(\exists m)(\overline{n}(m) > k)$, and finally $(\forall^{st} k)(\exists^{st} m)(\overline{n}(m) > k)$, by Σ_1-transfer. Tranfer for Π_2-formulas implies $(\forall k)(\exists m)(\overline{n}(m) > k)$. Thus far, the proof was very similar to the proof of theorem 1.138. However, the extra parameter ω in $\tau(n,\omega)$ now comes into play. Indeed, in this proof, $\overline{n}(m)$ is the largest $n' \leq m$ such that $(\forall n \leq n')(\tau(\sigma(n,m),m) < \tau(\sigma(n{+}1,m),m))$. Thus, for each M, there is an N such that for all $n \leq M$, we have $\tau(\sigma(n,N),N) < \tau(\sigma(n+1,N),N)$. However, we are interested in the behaviour of $\tau(\sigma(n,N),\omega)$, not in that of $\tau(\sigma(n,N),N)$. But since $\tau(n,\omega)$ is near-standard, we have $\tau(\sigma(n,N),N) \approx \tau(\sigma(n,N),\omega)$ for infinite N. Thus, $\tau(\sigma(n,N),N) < \tau(\sigma(n+1,N),N)$ implies $\tau(\sigma(n,N),\omega) \lessapprox \tau(\sigma(n+1,N),\omega)$, for infinite N and the theorem follows. □

In the previous theorems, the function σ is such that $\tau(\sigma(n))$ is monotone over $\mathbb{N}$. However, for each hyperfinite segment, we require a different function (or an equivalent Π_2-statement involving M and N). The reason is that in ERNA, there are numbers beyond $\mathbb{N}$, but not beyond the hypernaturals. Thus, a theory of Nonstandard Analysis in which there are always 'more infinite' numbers beyond any number, would be more elegant, as it can eliminate the Π_2-statement. Such is the topic of Chapter II.

Note that the function σ as in (1.46) is not near-standard, even for standard τ. Thus, bar transfer is useless in this context and this explains why we need a stronger principle, like Π_2-transfer, to prove the above theorems. The following theorem shows that Π_2-transfer is *exactly* what is needed.

1.141. THEOREM. *In* ERNA$^+$*, the Standard Subsequence principle is equivalent to* Π_2*-transfer.*

PROOF. By theorem 1.138, the inverse implication is immediate. For the forward implication, we use the unboundedness principle and corollary 2.59 from section 5.1. Let $f \in L^{st}$ be weakly increasing and such that $(\forall^{st} n)(\exists^{st} m)(f(m) > n)$ and $f(0) \geq 0$. By the Standard Subsequence principle, there is a function $\sigma : \mathbb{N} \to \mathbb{N}$ such that $f(\sigma(n,\omega))$ is strictly increasing or weakly decreasing over $\mathbb{N}$. By definition, $f(\sigma(n,\omega))$ cannot be weakly decreasing over $\mathbb{N}$ and hence it must be strictly increasing there. By the Standard Subsequence principle, for each M, there is an N such that $f(\sigma(n,N))$ is strictly increasing for $n \leq M$. In particular, this implies $f(\sigma(M,N)) > M$ and hence there follows $(\forall n)(\exists m)(f(m) > n)$. Thus, we have obtained the unboundedness principle, which implies Π_2-transfer by corollary 2.59 from section 5.1. □

A version of Ramsey's theorem is equivalent to ACA_0 over RCA_0 ([**46**, III.7.6]). It would be interesting to obtain an ERNA-version of this theorem. To define the infinite monochromatic set in Ramsey's theorem, we seem to need EXIT.

The following theorem is the ERNA$^+$-version of a well-known property of Archimedean fields.

1.142. THEOREM. *Let $\tau(n)$ be an internal sequence. Further, let a be a finite constant such that $\tau(n) \leq \tau(n+1) \leq a$ for all natural n. Then $\tau(n)$ is Cauchy.*

PROOF. If the assertion were false, there would exist some natural k_0 such that

$$(\forall^{st} N)(\exists^{st} n)(\exists m \leq n)\varphi(n,m,N). \tag{1.50}$$

where $\varphi(n,m,N)$ stands for

$$(m, n > N \wedge |\tau(n) - \tau(m)| \geq 1/k_0)$$

Denote the least $n \in \mathbb{N}$ such that $(\exists m \leq n)\varphi(n, m, N)$ by $f_1(N)$. Denote the least $m \leq n$ such that $\varphi(n, m, N)$ by $f_2(N, n)$. Finally we define $g(n) := f_2(h(n-1), h(n))$, where $h(n)$ is the n-th iteration of f_1 at zero. The latter function is available in ERNA becaus of example 1.37.

By construction, all intervals $\big]\tau(g(l)), \tau(h(l))\big]$ are disjoint and have length at least $1/k_0$. Therefore, $\sum_{n=2}^{n_0+2} \big|\tau(h(n)) - \tau(g(n))\big|$, with $n_0 = \lceil k_0|a - \tau(0)|\rceil$, would be larger than $a - \tau(0)$. This clearly is a contradiction, because a finite number of disjoint subintervals cannot have a total length exceeding that of the original interval.

By (1.50), the function $f_1(N)$ returns a natural number if $N \in \mathbb{N}$. Finally, the schema EXIT ensures that $h(n)$ is finite for finite input. $\square$

Using the previous theorem, we can prove the following version of the 'monotone convergence theorem' (see theorem 1.135). The proof takes place in $\text{ERNA}^+ + \Pi_1$-TRANS and, presumably, EXIT cannot be omitted.

1.143. Theorem (Monotone convergence principle). *A near-standard sequence $\tau(n)$ which is finitely bounded and weakly increasing for $n \in \mathbb{N}$, is convergent to $\tau(\omega)$. The terms of infinite index are infinitely close to each other.*

Proof. Let $\tau(n)$ be as stated. By the previous theorem, $\tau(n)$ is Cauchy and the theorem follows from theorem 1.98. $\square$

The following reversal is almost immediate.

1.144. Theorem. *In ERNA^+, the monotone convergence principle is equivalent to Π_1-TRANS.*

Proof. The inverse implication is immediate from the previous theorem. The forward direction is proved in the same way as in theorem 1.107. $\square$

Next, we consider the statement 'every bounded sequence of real numbers has a least upper bound' (see theorem 1.135). As we use theorem 1.143 in the proof of the following theorem, it takes place in $\text{ERNA}^+ + \Pi_1$-TRANS.

1.145. Theorem. *Every near-standard $\tau(n)$, finitely bounded on $\mathbb{N}$, has a least upper bound up to infinitesimals, i.e. there is a finite number L such that for all n, $\tau(n) \lessapprox L$ and for $K \ll L$, there is an $n_0 \in \mathbb{N}$ such that $K \ll \tau(n_0)$.*

Proof. Let $\tau(n)$ be as stated. Define $\rho(n)$ as $\max_{1 \leq i \leq n} |\tau(i)|$. Then $\rho(n)$ is weakly increasing, near-standard and finitely bounded on $\mathbb{N}$ and by theorem 1.143 this sequence converges to $\rho(\omega)$ and we have $\rho(\omega) \approx \rho(\omega')$ for infinite ω'. Defining $L := \rho(\omega)$, the previous implies $\tau(n) \lessapprox L$ for all n and applying bar transfer on the boundedness condition of τ yields that L is finite. By definition, $K \ll L$ implies that there is an $m \leq \omega$ such that $K \ll \tau(m)$. Using bar tranfer, it is easy to prove the existence of the number n_0 of the theorem. $\square$

As expected, the previous theorem is also equivalent to Π_1-TRANS.

1.146. Theorem. *In ERNA^+, theorem 1.145 is equivalent to Π_1-TRANS.*

Proof. The inverse implication is immediate from the previous theorem. For the forward implication, assume theorem 1.145 and let τ be as stated there. Let φ be as stated in Π_1-TRANS and assume $\varphi(n)$ for all $n \in \mathbb{N}$. Define $\sigma(n)$ as in (1.41). By assumption, we have $\sigma(n) = \tau(n)$ for $n \in \mathbb{N}$ and hence $\sigma(n)$ satifies the conditions of theorem 1.145. Thus, there is a finite number L such that $\sigma(n) \lessapprox L$ for all n and this fact excludes the possibility that $\sigma(m) = m$ for infinite m. Hence, the case $\sigma(n) = n$ never occurs and we have $\sigma(n) = \tau(n)$ for all n. This implies $\varphi(n)$ for all n and we have obtained Π_1-TRANS. □

Let $\mathbb{T}$ be theorem 1.145 with the occurence of 'on $\mathbb{N}$' on its first line omitted. Interestingly, the weaker version $\mathbb{T}$ is still equivalent to Π_1-TRANS. However, we need a more subtle argument, as $\sigma(n)$ from (1.41) is only provably bounded on $\mathbb{N}$.

1.147. Theorem. *In* ERNA$^+$*,* $\mathbb{T}$ *is equivalent to* Π_1-TRANS.

Proof. We only need to prove that $\mathbb{T}$ implies Π_1-TRANS. Thus, let τ be as in $\mathbb{T}$ and let φ be as in Π_1-TRANS. Note that we may assume that $0 \leq \tau(n) \leq 1$ for all n. Assume $\varphi(n)$ for all $n \in \mathbb{N}$ and suppose there is an m_0 such that $\neg\varphi(m_0)$ and let m_1 be the least of these. By assumption, the number m_1 is infinite. Let $q > 0$ be a rational such that $\tau(n) \leq q$ for $n < m_1$. Define

$$\sigma'(n) := \begin{cases} \tau(n) & (\forall m \leq n)\varphi(m) \\ \tau(n) + 2q & \text{otherwise} \end{cases}. \tag{1.51}$$

Then $\sigma'(n)$ is bounded *everywhere* and hence $\mathbb{T}$ applies. As $\neg\varphi(m_1)$, we have $\sigma(m_1) = \tau(m_1) + 2q$ and hence L is at least $2q$. By the leastness of L, there is a term $\sigma'(n_0)$ of finite index n_0 between L and $L - q$. In particular, $\sigma'(n_0) > q$ and, by assumption, $\sigma'(n_0) = \tau(n_0)$. However, by the definition of q, there are no terms $\tau(n)$ of finite index above q. This is a contradiction and thus the number m_0 cannot exist. Hence, $\varphi(n)$ must hold for all n and Π_1-TRANS follows. □

This theorem is interesting, as the omission of 'on $\mathbb{N}$' in theorem 1.145 changes the latter fundamentally. Indeed, in its original form, theorem 1.145 has the structure '$A \to B$', where A is a property of $\mathbb{N}$ and B is a property of *all* hypernaturals. Thus, it is not inconceivable that the schema consisting of the implications $A \to B$ may be equivalent to Π_1-TRANS, as this transfer principle is the archetype of the structure '$A \to B$' mentioned before. However, by omiting 'on $\mathbb{N}$' in theorem 1.145, this structure is lost, but the resulting theorem $\mathbb{T}$ is still equivalent to Π_1-TRANS.

Let $\mathbb{S}$ be the Supremum Principle formulated in theorem 1.100. The latter is ERNA's version of Dedekind completeness. As Cauchy completeness is classically equivalent to Dedekind completeness (in the sense of [**46**, III.2.2]), we expect $\mathbb{S}$ to be equivalent to Π_1-TRANS, by theorem 1.107. However, $\mathbb{S}$ is very similar to $\mathbb{T}$ from the previous paragraph. Indeed, let $\mathbb{S}'$ be $\mathbb{S}$ with item (i) replaced by '$\varphi(x)$ holds for no *rational* $x > b$'. Then it is fairly obvious that $\mathbb{S}'$ equivalent to Π_1-TRANS, as the implication $(\forall^{st} x > b)\varphi(x) \to (\forall x > b)\varphi(x)$ is explicitly embedded in $\mathbb{S}'$. However, $\mathbb{S}$ does not have this form and neither does $\mathbb{T}$, as discussed previously. Hence, it is not obvious that $\mathbb{S}$ and Π_1-TRANS are equivalent. The following theorem asserts the equivalence, but the proof is subtle and would not have been discovered without studying $\mathbb{T}$ first.

1.148. Theorem. *In* ERNA*, the Supremum Principle is equivalent to* Π_1-TRANS.

Proof. The inverse implication is immediate from theorem 1.100. For the forward implication, let φ_1 and b be such that $\varphi_1(x)$ holds for no $x > b$ and that there are a of arbitrary weight such that $\varphi_1(a)$ holds, with $a, b \gg 0$ and b rational. Let φ_2 be as in Π_1-TRANS and assume $\varphi_2(n)$ holds for $n \in \mathbb{N}$. Now suppose that there is an m_0 such that $\neg\varphi_2(m_0)$ and let $a_0 \gg 0$ be such that $\varphi_1(a_0)$ and $\|a_0\| > \|b\|m_0$. It is an easy verification that the following formula can be defined in ERNA:

$$\psi(x) \equiv \begin{cases} \varphi_1(x) & (\forall n \leq \|x\|)\varphi_2(n) \\ \varphi_1(x-b) & \text{otherwise} \end{cases}.$$

By definition, we have $\neg\psi(x)$ for $x > 2b$ and $\psi(a_0 + b)$. By definition, the number β provided by the Supremum Principle is at least $a_0 + b$. Let $\varepsilon \gg 0$ be such that $\beta - \varepsilon > b$. By the Supremum Principle, there is a rational x_0 such that $\psi(x_0)$ and $x_0 > \beta - \varepsilon$. Hence, $x_0 > b$ and since x_0 is rational, we have $\psi(x_0) \equiv \varphi_1(x_0)$, by definition. But $\varphi_1(x)$ does not hold for $x > b$ and we have obtained a contradiction. Thus, the number m_0 cannot exist and the formula $\varphi_2(n)$ holds for all n. This implies Π_1-TRANS and we are done. □

Note that the implication $(\exists x < b)\varphi(x) \to (\exists^{st} x < b)\varphi(x)$ is embedded in $\mathbb{S}$ thanks to items (ii) and (iv) of this schema. Thus, Π_1-TRANS is embedded in $\mathbb{S}$ in the form of the equivalent schema Σ_1-TRANS. Moreover, if $\mathbb{S}_n$ is $\mathbb{S}$ with $\varphi \in \Pi_{n-1}$, then the previous theorem implies that $\mathbb{S}_1$ is equivalent to Π_1-TRANS. In the same way, it is clear that $\mathbb{S}_2$ ($\mathbb{S}_3$) is equivalent to Π_2-TRANS (Π_3-TRANS) because of items (ii) and (iv) in $\mathbb{S}_2$ ($\mathbb{S}_3$). To prove the same for $\mathbb{S}_n$ with $n > 3$, use theorem 3.8.

Finally, we can prove ERNA's version of the Bolzano-Weierstraß theorem.

1.149. Theorem (Internal Bolzano-Weierstraß). *For every $\tau(n)$, finitely bounded on $\mathbb{N}$, there is a function $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\tau(\sigma(n))$ converges to some term on infinite index. The terms with small enough infinite index are infinitely close.*

Proof. Immediate from corollary 1.99 and theorems 1.137 and 1.142. □

The following theorem shows that, if Π_2-transfer is available, we can extend the Cauchy property of $\tau(\sigma(n))$ to arbitrarily long initial segments. Thus, it is proved in $\text{ERNA}^+ + \Pi_2$-TRANS.

1.150. Theorem (Standard Bolzano-Weierstraß). *For every $\tau(n) \in L^{st}$, finitely bounded on $\mathbb{N}$, there is a function $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\tau(\sigma(n,\omega))$ converges to some $\tau(\sigma(m_0,\omega))$ with m_0 infinite. Also, for each M, there is an N such that all terms $\tau(\sigma(n,N))$ with infinite index $n \leq M$ are infinitely close.*

Proof. Let τ be as stated. Let $\sigma(n,\omega)$ be the function provided by the Standard Subsequence principle. By theorem 1.142, the sequence $\tau(\sigma(n,\omega))$ has the Cauchy property, i.e. for all $k \in \mathbb{N}$, there is an $N \in \mathbb{N}$ such that

$$(\forall^{st} m, m' > N)\big[|\tau(\sigma(m,\omega)) - \tau(\sigma(m',\omega))| < 1/k\big]. \tag{1.52}$$

By overflow, we obtain $\overline{m}(k,\omega)$ which is infinite for all $k \in \mathbb{N}$. Thus, we have $(\forall^{st} k)(\overline{m}(k,\omega) > k)$. By overflow, there follows $(\forall k \leq \overline{k}(\omega))(\overline{m}(k,\omega) > k)$, where $\overline{k}(\omega)$ is infinite. Note that $\overline{k}(m)$ is infinite for infinite m. Define $\overline{m}(l)$ as the least of all $\overline{m}(k,l)$ with $k \leq \overline{k}(l)$. Then $\overline{m}(\omega)$ is infinite and in the same way as in theorem 1.138, we have $(\forall^{st} l)(\exists^{st} l')(\overline{m}(l') > l)$. By Π_2-transfer, this yields $(\forall l)(\exists l')(\overline{m}(l') > l)$.

Now let M be arbitrary and choose L such that $\overline{m}(L) > M$. By the Standard Subsequence Principle, $\tau(\sigma(n,L))$ is monotone on $\mathbb{N}$. By theorem 1.142, the sequence

$\tau(\sigma(n,L))$ has the Cauchy property, i.e. for all $k \in \mathbb{N}$, there is an $N \in \mathbb{N}$ such that

$$(\forall^{st} m, m' > N)\big[|\tau(\sigma(m,L)) - \tau(\sigma(m',L))| < 1/k\big]. \tag{1.53}$$

By overflow, we obtain $\overline{m}(k,L)$, which is infinite for $k \in \mathbb{N}$. By definition, we have $\overline{m}(k,L) \geq \overline{m}(L)$, for $k \in \mathbb{N}$, which implies

$$(\forall m, m' \leq M)\big[m, m' > N \to |\tau(\sigma(m,L)) - \tau(\sigma(m',L))| < 1/k\big]. \tag{1.54}$$

Thus, all terms of infinite index $m \leq M$ are infinitely close. □

1.151. Corollary. *The theorem also holds for near-standard sequences.*

Proof. The proof of the theorem can be copied and we obtain (1.54) with $\tau(\sigma(n,L),L)$ instead of $\tau(\sigma(n,L))$. By definition, $\tau(\sigma(n,L),L)$ is infinitely close to $\tau(\sigma(n,L),\omega)$ if L is infinite. Thus, the corollary follows. □

The following corollary is a reformulation of the theorem and corollary in more intuitive wording.

1.152. Corollary. *For any near-standard sequence, finitely bounded on $\mathbb{N}$, and any infinite number M, there's a convergent subsequence with limit of index M.*

By our earlier results, it is clear that theorem 1.150 implies Π_1-TRANS. However, we used Π_2-TRANS in the proof of this theorem, so we do not have equivalence between the latter and Π_1-TRANS. Also, to avoid trivialities, the convergent subsequence in theorem 1.150 should in general not be constant beyond some infinite index. Hence the following corollary.

1.153. Corollary (BW). *For every M and $\tau(n) \in L^{st}$, finitely bounded on $\mathbb{N}$, there is a function $\sigma : \mathbb{N} \to \mathbb{N}$ and an N such that all terms $\tau(\sigma(m,N))$ with infinite index $m \leq M$ are infinitely close limits of $\tau(\sigma(n,N))$. These limits are non-identical if $\tau(\sigma(n,\omega))$ is strictly increasing on $\mathbb{N}$.*

Proof. The corollary follows from the theorem, except for the last sentence. In case $\tau(\sigma(n,\omega))$ is strictly increasing, replace (1.52) with

$$(\forall^{st} m, m' > N)\big[m \neq m' \to 0 < |\tau(\sigma(m,\omega)) - \tau(\sigma(m',\omega))| < 1/k\big].$$

The rest of the proof is identical. □

1.154. Theorem (Nonstandard Bolzano-Weierstraß). *For each standard sequence, finitely bounded on $\mathbb{N}$, and each infinite M, there is a convergent subsequence with limit of index M and with terms differing less than $1/M$.*

Proof. The first part of the theorem, up to 'with limit of index M' is immediate from corollary 1.152. For the remaining part, let the sequence τ be as stated. Let $\sigma(n,\omega)$ be the function provided by the Standard Subsequence principle. By theorem 1.142, the sequence $\tau(\sigma(n,\omega))$ has the Cauchy property, i.e. for all $k \in \mathbb{N}$, there is an $N \in \mathbb{N}$ such that

$$(\forall^{st} m, m' > N)\big[|\tau(\sigma(m,\omega)) - \tau(\sigma(m',\omega))| < 1/k\big].$$

By overflow, we obtain $\overline{m}(k,\omega)$, which is infinite for all $k \in \mathbb{N}$. Thus, we have

$$(\forall^{st} k)(\forall m, m' \in [\overline{m}(k,\omega)/2, \overline{m}(k,\omega)])\big[|\tau(\sigma(m,\omega)) - \tau(\sigma(m',\omega))| < 1/k\big].$$

Overflow yields the term $\overline{k}(\omega)$. In the same way as in the previous proofs, we obtain $(\forall^{st} l)(\exists^{st} l')\overline{k}(l') > l$ and, by Π_2-transfer, $(\forall l)(\exists l')\overline{k}(l') > l$.

Now let K be arbitrary and choose L such that $\overline{k}(L) > K$. By the Standard Subsequence Principle, $\tau(\sigma(n, L))$ is monotone on $\mathbb{N}$. By theorem 1.142, the sequence $\tau(\sigma(n, L))$ has the Cauchy property, i.e. for all $k \in \mathbb{N}$, there is an $N \in \mathbb{N}$ such that

$$(\forall^{st} m, m' > N)\big[|\tau(\sigma(m, L)) - \tau(\sigma(m', L))| < 1/k\big]. \tag{1.55}$$

By overflow, we obtain $\overline{m}(k, L)$, which is infinite for $k \in \mathbb{N}$. Thus, we have

$$(\forall^{st} k)(\forall m, m' \in [\overline{m}(k, L)/2, \overline{m}(k, L)])\big[|\tau(\sigma(m, L)) - \tau(\sigma(m', L))| < 1/k\big].$$

By overflow, we obtain $\overline{k}(L)$, which is at least K. Thus, terms of the sequence $\tau(\sigma(m, L))$ of index $m, m' \in [\overline{m}(\overline{k}(L), L)/2, \overline{m}(\overline{k}(L)), L)]$ differ at most $1/K$. □

In the previous proof, we did *not* obtain the Cauchy property (see 1.36) with unbounded quantifiers $(\forall k)$, $(\exists N)$ and $(\forall n, m)$. We believe that the latter requires Π_3-TRANS.

In the following theorem, we prove that one of ERNA's versions of the Bolzano-Weierstraß theorem is equivalent to Π_2-TRANS. Many variations are possible.

1.155. Theorem. *In* ERNA$^+$, BW *is equivalent to* Π_2-TRANS.

Proof. The inverse implication is immediate from corollary 1.153. For the forward implication, we use the unboundedness principle and corollary 2.59 from section 5.1. Let $f \in L^{st}$ be weakly increasing with $(\forall^{st} n)(\exists^{st} m)(f(m) > n)$. Define $\tau(n)$ as $1 - \frac{1}{f(n)}$. By BW, for every M, there is an N such that $\tau(\sigma(m, N)) \approx \tau(\sigma(m', N))$ for infinite $m, m' \leq M$. Note that for $m \neq m'$, the terms are not identical. As f is weakly increasing, there must hold that $f(n)$ becomes arbitrarily large, i.e. $(\forall n)(\exists m)(f(m) > n)$. By corollary 2.59, Π_2-TRANS follows and we are done. □

4.3.2. *External affairs.* In the previous paragraph, we worked in ERNA$^+$. However, since EXIT is not needed for the Reverse Mathematics of ERNA+Π_1-TRANS, the theory ERNA$^+$ is not a suitable base theory. In this paragraph, we study equivalent formulations of EXIT and related schemas, like the following.

1.156. Axiom schema (ATOM). *For every arithmetical f, if $f(n)$ is infinite for $n \in \mathbb{N}$, then there is a least number with this property.*

1.157. Definition. The class Π_0^{st} consists of all bounded standard formulas. For $n \geq 1$, Π_n^{st}, is the class of Π_n-formulas with standard quantifiers $(\forall^{st} n)$ and $(\exists^{st} m)$ instead of the usual quantifiers $(\forall n)$ and $(\exists m)$.

1.158. Axiom schema (Π_n^{st}-MIN). *For $\varphi \in \Pi_n^{st}$, if there is an $m \in \mathbb{N}$ such that $\varphi(m)$, then there is a least number with this property.*

In [**1**], Avigad discusses Π_2^{st}-MIN and related schemas in the context of Reverse Mathematics. We answer some of the questions from [**1**, §6] in paragraph 4.3.3.

1.159. Theorem. *In* ERNA + Π_1-TRANS*,* ATOM *is equivalent to* Π_2^{st}-MIN*.*

Proof. For the inverse implication, assume Π_2^{st}-MIN and let f be as in ATOM. Let $f(n, m)$ be $f(n)$ with all occurrences of ω replaced with the free variable m. Then $f(n)$ is $f(n, \omega)$. Now assume $f(n_0, \omega)$ is infinite for $n_0 \in \mathbb{N}$. This implies $(\forall^{st} k)(f(n_0, \omega) > k)$ and also $(\forall^{st} k)(\exists N)(f(n_0, N) > k)$, and, by Σ_1-transfer, $(\forall^{st} k)(\exists^{st} N)(f(n_0, N) > k)$. Let $n_1 \in \mathbb{N}$ be the least n such that $(\forall^{st} k)(\exists^{st} N)(f(n, N) > k)$, as provided by Π_2^{st}-MIN. By leastness, we have $(\exists^{st} k)(\forall^{st} N)(f(n_1-$

$1, N) \le k)$ and applying Π_1-transfer to the innermost universal formula yields $(\exists^{st} k)(\forall N)(f(n_1 - 1, N) \le k)$. In particular, we have $(\exists^{st} k)(f(n_1 - 1, \omega) \le k)$. Hence, $f(n_1 - 1, \omega)$ is finite and the same holds for all $n < n_1$. As $f(n, N)$ is weakly increasing in N, $f(n_1, \omega)$ is infinite. Thus, n_1 is the least n such that $f(n, \omega)$ is infinite, which implies ATOM.

For the forward implication, assume ATOM and let $(\forall^{st} n)(\exists^{st} m)\varphi(n, m, k)$ be in Π_2^{st}. In ERNA $+ \Pi_1$-TRANS, this formula is equivalent to

$$(\forall n \le \overline{n}(\omega, k))(\exists m \le \omega)\varphi(n, m, k) \wedge \overline{n}(k, \omega) \text{ is infinite.} \tag{1.56}$$

Let $\psi(k)$ be the first part of the conjunction. By corollary 1.46, we may treat $\psi(k)$ as quantifier-free. Let $T_\psi(k)$ be the function obtained from theorem 1.45. Then (1.56) is equivalent to

$$\big[T_\psi(k) \times \overline{n}(k, \omega)\big] \text{ is infinite.} \tag{1.57}$$

Thus, we see that if there is a $k \in \mathbb{N}$ such that $(\forall^{st} n)(\exists^{st} m)\varphi(n, m, k)$, the schema ATOM applied to (1.57) gives us the least of these. □

An analogous result exists for Π_3^{st}-MIN, see theorem 2.68. Note that Π_2^{st}-MIN does not involve ω and $\approx$ (except in the quantifiers), whereas ATOM does. By contraposition, ATOM is equivalent to the following schema.

1.160. Axiom schema (ATI). *For every arithmetical f, if $f(0)$ is finite and '$f(n)$ is finite' implies '$f(n+1)$ is finite', for all $\mathbb{N}$; then $f(n)$ is finite for $n \in \mathbb{N}$.*

Note that ATI is the formalization of the physical intuition 'no finite operation iterated finitely many times can reach the infinite'. By contrast, Π_2^{st}-MIN is a purely logical schema. As ATOM implies EXIT, we have the following theorem.

1.161. Theorem. *In* ERNA*, the following are equivalent.*
1. Π_1-TRANS $+ \Pi_2^{st}$-MIN .
2. *The monotone convergence theorem plus* ATI.
3. $\mathbb{T}$ + ATI.
4. *Theorem 1.145 plus* ATI.

In ERNA*, the following are equivalent.*

5. Π_2-TRANS $+ \Pi_2^{st}$-MIN.
6. *The Standard Subsequence principle plus* ATI.
7. BW *plus* ATI.

In line with the theme of Reverse Mathematics, we have obtained equivalences between pairs of ordinary mathematical statements and pairs of logical statements. Note that none of the equivalences follows (in ERNA) if we omit one of the members of a pair. Thus, the members of the pairs are somehow 'interlaced'.

In section 4.2, we speculated on the connection between the Reverse Mathematics of ERNA $+ \Pi_1$-TRANS and Constructive Reverse Mathematics. Recently, Hajime Ishihara showed that the statement 'Every bounded increasing sequence of reals is a Cauchy sequence' is equivalent to a principle slightly stronger than Σ_1-PEM. Also, Ishihara showed that Π_1^0-AC_{00} implies 'Every Cauchy sequence has a modulus'. Compare this with item 2 of theorem 1.161, principle 1.127 and theorem 1.134.

With an eye on future research, we mention that the Ascoli-Arzela theorem is equivalent to the Bolzano-Weierstraß theorem ([**46**, III.2.9]).

4.3.3. *The theory* $\mathbf{NERA}^\omega$. The theory NERA^ω is essentially the higher type extension of ERNA and Avigad develops very elementary calculus in this theory (see [**1**] for details). Many definitions are similar to ours, compare e.g. the definition of $r \leq_{\mathbb{R}} s$ with our definition of $r \lessapprox s$, and Π_1-transfer is even implicitly built into the definition of a formula 'which respects equality of reals'. However, instead of working with the hyperrationals and obtaining results 'up to infinitesimals', a model of the real numbers with equality $=_{\mathbb{R}}$ is obtained by taking the quotient of the hyperrationals with the relation $\approx$. This alternative approach yields the peculiar result that all functions $f : \mathbb{R} \to \mathbb{R}$ are continuous. Although this is well-known to be correct in the intuitionistic setting, this poses problems for the further development of classical calculus in NERA^ω. Furthermore, in the conclusion of [**1**], Avigad list three important questions with respect to Nonstandard Analysis.

(1) Is there a better way [than '$=_{\mathbb{R}}$'] to treat equality?
(2) What is required to formalize various parts of analysis?
(3) Do nonstandard theories provide a useful approach?

In our opinion, the following answers are given in this dissertation.

(1) Simply replace equality with '$\approx$'. By theorem 1.3, many well-known theorems and equivalences hold 'up to infinitesimals'.
(2) When it comes to analysis in WKL_0, the schema Π_1-TRANS is exactly what is needed, if an infinitesimal error is allowed.
(3) In light of theorem 1.3, the answer to this question can only be positive.

Moreover, in [**1**, §6 and Appendix A]), Avigad discusses the role of arithmetical axiom schemas like Π_2^{st}-MIN. In particular, he poses the question whether such schemas are natural and if they play a significant role in Reverse Mathematics. Theorem 1.161 suggests that the answer to these questions is also positive.

4.3.4. *Second-order arithmetic.* In this section we discuss an extension of ERNA+ Π_1-TRANS to the framework of second-order arithmetic. We use Yokoyama's notation from [**57**]. Our conclusion is that a full second-order version of ERNA + Π_1-TRANS has first-order strength of Peano arithmetic, which goes far beyond $I\Delta_0 + \exp$. This is one of the reasons why we choose the first-order theory ERNA, rather than a higher-order extension.

The second-order theory $\mathbb{A}$, in the language $\mathcal{L}_2^*$, consists of the following axioms.

- Induction for bounded $\mathcal{L}_2^*$-formulas.
- Comprehension for bounded $\mathcal{L}_2^*$-formulas.
- Standard Part: $(\forall X^*)(\exists Y^s)(\forall x^s)(\check{x}^s \in X^* \leftrightarrow x^s \in Y^s)$.
- Σ_1-transfer: $(\forall x^s, X^s)(\varphi(x^s, X^s)^s \leftrightarrow \varphi(\check{x}^s, \check{X}^s)^*)$ $(\varphi \in \Sigma_1)$.

The following theorem shows that $\mathbb{A}$ implies arithmetical comprehension.

1.162. Theorem. *The theory* $\mathbb{A}$ *proves the existence of every set* $\{n \in \mathbb{N} | \psi(n)\}$, *with* ψ *arithmetical and standard.*

Proof. First, we treat the case $\psi \in \Sigma_1$. Let $\varphi(x, y, X)$ be a bounded $\mathcal{L}_2$-formula with set parameter $X \subset \mathbb{N}$. For readability, we surpress possible standard number parameters. By Δ_0-comprehension, there is a set Y^* such that $y^* \in Y^* \leftrightarrow (\exists x^* \leq \omega)\varphi(x^*, y^*, \check{X}^s)^*$. By the standard part axiom, there is a standard set Z^s with the same standard elements as Y^*. Thus, for $y^s \in \mathbb{N}$, we have $y^s \in Z^s \leftrightarrow (\exists x^* \leq \omega)\varphi(x^*, y^s, \check{X}^s)$. By the Σ_1-transfer principle, the latter is equivalent to $(\exists x^s)\varphi(x^s, y^s, X^s)$ and this case is done.

The general case now follows easily. Indeed, by the previous case, a standard Σ_1 or Π_1-formula can be reduced to an equivalent standard quantifier-free formula. Applying this reduction inductively, we see that a Σ_n (Π_n) formula is equivalent to a Σ_1 or Π_1-formula. Thus, the general case follows from the particular case for $n = 1$. $\square$

Note that, in the same way as in the proof, Σ_1-ACA implies arithmetical comprehension (see [**46**, III.1.4]).

Thus, theorem 1.162 suggests that the Reverse Mathematics of the theory ERNA + Π_1-TRANS cannot be directly generalized to the usual second-order framework. A possible solution is omitting (part of) the Standard Part axiom. Note that this would be a step towards 'internal' nonstandard analysis (see Chapter III).

Using nonstandard analysis, Keisler has developed an axiomatization of the Big Five of Reverse Mathematics by showing that nonstandard numbers can code subsets of $\mathbb{N}$ ([**32**]). However, theorems 1.3 and 1.162 suggest that the nonstandard framework yields a more subtle picture of Reverse Mathematics than the second-order framework. From a utilitarian point of view, this implies that nonstandard numbers are 'more real' than subsets of $\mathbb{N}$.

1.163. Acknowledgement. I wish to thank Karel Hrbacek, Chris Impens, Hajime Ishihara, Kazuyuki Tanaka, Andreas Weiermann and Keita Yokoyama for their valuable advice with regard to this chapter. Special thanks goes to the anonymous referee who pointed out errors in an earlier draft of this chapter.

CHAPTER II

Beyond ε-δ: Relative infinitesimals and ERNA

> Logicians are perhaps more
> philosophers than mathematicians.
>
> *LC 2007, Wrocław, Poland*
> Andrzej Grzegorczyk

1. Introduction

The theories ERNA and NQA$^+$ are said to 'provide a foundation that is close to mathematical practice characteristic of theoretical Physics', according to Chuaqui, Sommer and Suppes. In order to achieve this goal, the systems satisfy the following three conditions, listed in [**11**]:

(i) The formulation of the axioms is essentially a free-variable one with no use of quantifiers.
(ii) We use infinitesimals in an elementary way drawn from Nonstandard Analysis, but the account here is axiomatically self-contained and deliberately elementary in spirit.
(iii) Theorems are left only in approximate form; that is, strict equalities and inequalities are replaced by approximate equalities and inequalities. In particular, we use neither the notion of standard function nor the standard-part function.

It is also mentioned in [**11**], that another standard practice of Physics, namely the use of physically intuitive but mathematically unsound reasoning, is not reflected in the system.

By limiting the strength of the systems according to (i)-(iii), the consistency of ERNA can be proved in PRA, using Herbrand's theorem in the form described in the previous chapter. In this respect, the item (i) is not merely a technicality to suit Herbrand's theorem: the quantifier-free axioms reflect the absence[1] of existential quantifiers in Physics. As all ε-δ definitions of basic analysis are equivalent to universal nonstandard formulas, it indeed seems plausible that one can develop calculus inside ERNA in a quantifier-free way, particularly, without the use of ε-δ-statements. However, we discuss two compelling arguments why such a development is impossible.

First, as exemplified by item (iii), NQA$^+$ has no 'standard-part' function 'st', which maps every finite number x to the unique standard number y such that $x \approx y$. Thus,

[1] Patrick Suppes (Stanford University) is said to offer substantial monetary rewards for examples to the contrary.

nonstandard objects like integrals and derivatives are only defined 'up to infinitesimals'. This leads to problems when trying to prove e.g. the fundamental theorems of calculus, which express that differentiation and integration cancel each other out. Indeed, in [**11**, Theorem 8.3], Chuaqui and Suppes prove the first fundamental theorem of calculus, using the previously proved corollary 7.4. The latter states that differentiation and integration cancel each other out *on the condition* that the mesh du of the hyperfinite Riemann sum of the integral and the infinitesimal y used in the derivative satisfy $du/y \approx 0$. Thus, for every y, there is a du such that for all meshes $dv \leq du$ the corresponding integral and derivative cancel each other out. The definition of the Riemann integral ([**11**, Axiom 18]) absorbs this problem, but the former is quite complicated as a consequence. Also, it does not change the fact that ε-δ-statements occur, be it swept under the proverbial nonstandard carpet. Similary, ERNA only proves a version of the first fundamental theorem and of Peano's existence theorem with a condition similar to $du/y \approx 0$, *contrary* to Sommer and Suppes' claim in [**50**]. (see theorems 1.94 and 1.96). Thus, ERNA and NQA^+ cannot develop basic analysis without invoking ε-δ statements.

Second, we consider to what extent that classical Nonstandard Analysis is actually free of ε-δ-statements. For all functions in the standard language, the well-known classical ε-δ definitions of continuity or Riemann integration, which are Π_3, can be replaced by universal nonstandard formulas (see e.g. [**48**, p. 70]). Given that even most mathematicians find it difficult to work with a formula involving more than two quantifier alternations, this is a great virtue. Indeed, using the nonstandard method greatly reduces the sometimes tedious 'epsilon management' when working with several ε-δ statements, see [**54**]. Yet, Nonstandard Analysis is not completely free of ε-δ statements. For instance, consider the function $\delta(x) = \frac{1}{\pi}\frac{\varepsilon}{\varepsilon^2+x^2}$, with $\varepsilon \approx 0$ and let $f(x)$ be a standard C^∞ function with compact support. Calculating the (nonstandard) Riemann integral of $\delta(x) \times f(x)$ yields $f(0)$. Hence, $\delta(x)$ is a nonstandard version of the Dirac delta. However, not every Riemann sum with infinitesimal mesh is infinitely close to the Riemann integral: the mesh has to be *small enough* (compared to ε). Moreover, $\delta(x) \approx \delta(y)$ is not true for all $x \approx y$, only for x and y *close enough.* In general, most functions which are not in the standard language do *not* have an elegant universal definition of continuity or integration and we have to resort to ε-δ statements. Thus, Nonstandard Analysis only partially removes the ε-δ formalism.

These two arguments show that the 'regular' nonstandard framework does not allow us to develop basic analysis in a quantifier-free way in weak theories of arithmetic. Moreover, for treating more advanced analysis, like the Dirac delta, prevalent in Physics, we would have to resort to ε-δ-statements anyway. Inspired by Hrbacek's 'stratified analysis' (see [**22**] and [**23**]), we introduce a weak theory of arithmetic, called $\text{ERNA}^{\mathbb{A}}$, which will allow us to develop analysis in a quantifier-free way. To this end, the theory $\text{ERNA}^{\mathbb{A}}$ has a multitude of sets of infinite numbers instead of the usual dichotomy of one set of finite numbers O, complemented with one set of infinite numbers Ω. Indeed, in $\text{ERNA}^{\mathbb{A}}$ there is a linear ordering $(\mathbb{A}, \preceq)$ with least number $\mathbf{0}$, such that for all nonzero $\alpha, \beta \in \mathbb{A}$, the infinite number ω_α is finite *compared to* ω_β for $\beta \succ \alpha$. Hence, there are many 'degrees' or 'levels' of infinity and the least number $\mathbf{0}$ in the ordering $(\mathbb{A}, \preceq)$ corresponds to the standard level. It should be noted that the first nonstandard set theory involving different levels of infinity was introduced by Péraire in [**40**]. Another approach was developed by Gordon in [**20**].

In the second section, we describe ERNA$^{\mathbb{A}}$ and its fundamental features and in the third section, we prove the consistency of ERNA$^{\mathbb{A}}$ inside PRA. Though important in its own right, in particular for 'strict' finitism (see [**51**]), we not only wish to do quantifier-free analysis in ERNA$^{\mathbb{A}}$ (see 4), but also study its metamathematics. Thus, in the section 3, we introduce the 'Stratified Transfer Principle', which expresses that a true formula should hold at all levels (see also [**22**]). Stratified Transfer equally applies to external formulas and is thus very different from transfer principles in regular nonstandard arithmetic. In the same section, we also introduce various transfer principles for ERNA$^{\mathbb{A}}$, which are based on transfer principles for ERNA. It turns out that the 'regular' transfer principle for Π_3-formulas is equivalent to the Stratified Transfer Principle, which is remarkable, given the fundamental difference in scope between both. This ERNA$^{\mathbb{A}}$-contribution to Reverse Mathematics has implications for ERNA and thus, in the section 5, we argue that techniques, ideas and even proofs carry over between ERNA and ERNA$^{\mathbb{A}}$. In particular, we prove several ERNA-theorems which would not have been discovered without studying ERNA$^{\mathbb{A}}$. Thus, there is an intimate connection between the stratified and classical nonstandard framework. Indeed, the former is a refinement of the latter, not a departure from it.

2. ERNA$^{\mathbb{A}}$, the system

In this section, we describe ERNA$^{\mathbb{A}}$ and some of its fundamental features.

2.1. Language and axioms.

2.1.1. *The language of* ERNA$^{\mathbb{A}}$. Let $(\mathbb{A}, \preceq)$ be a fixed linear order with least element $\mathbf{0}$, e.g. $(\mathbb{N}, \leq)$ or $(\mathbb{Q}^+, \leq)$. For brevity, we write '$\alpha \prec \beta$' instead of '$\alpha \preceq \beta \wedge \alpha \neq \beta$'.

2.1. Definition. The language L of ERNA$^{\mathbb{A}}$ includes ERNA's, minus the symbols 'ω', 'ε' and '$\approx$'. Additionally, it contains, for every nonzero $\alpha \in \mathbb{A}$, two constants 'ω_α' and 'ε_α' and, for every $\alpha \in \mathbb{A}$, a binary predicate '$\approx_\alpha$'.

The set $\mathbb{A}$ and the predicate $\preceq$ are not part of the language of ERNA$^{\mathbb{A}}$. However, we shall sometimes informally refer to them in theorems and definitions. Note that there are no constants $\omega_{\mathbf{0}}$ and $\varepsilon_{\mathbf{0}}$ in L.

2.2. Definition. For all $\alpha \in \mathbb{A}$, the formula '$x \approx_\alpha 0$' is read 'x is α-infinitesimal', 'x is α-infinite' stands for '$x \neq 0 \wedge 1/x \approx_\alpha 0$'; '$x$ is α-finite' stands for 'x is not α-infinite'; 'x is α-natural' stands for 'x is hypernatural and α-finite'.

2.3. Definition. If L is the language of ERNA$^{\mathbb{A}}$, then $L^{\alpha\text{-}st}$, the α-standard language of ERNA$^{\mathbb{A}}$, is L without $\approx_\beta$ for all $\beta \in \mathbb{A}$ and without ω_β and ε_β for $\beta \succ \alpha$.

For $\alpha = \mathbf{0}$, we usually drop the addition '$\mathbf{0}$'. For instance, we write 'natural' instead of '$\mathbf{0}$-natural' and '$\approx$' instead of '$\approx_{\mathbf{0}}$'. Note that in this way, $L^{\mathbf{0}\text{-}st}$ is L^{st}, the *standard* language of ERNA$^{\mathbb{A}}$.

2.4. Definition. A term or formula is called *internal* if it does not involve $\approx_\alpha$ for any $\alpha \in \mathbb{A}$; if it does, it is called *external*.

2.1.2. *The axioms of* ERNA$^{\mathbb{A}}$. The axioms of ERNA$^{\mathbb{A}}$ include ERNA's, minus axiom 1.10.4 (Hypernaturals), axiom set 1.14 (Infinitesimals) and axiom set 1.40 (External minimum). Additionally, ERNA$^{\mathbb{A}}$ contains the following axiom set.

2.5. Axiom set (Infinitesimals).

(1) *If x and y are α-infinitesimal, so are $x+y$ and $x \times y$.*
(2) *If x is α-infinitesimal and y is α-finite, xy is α-infinitesimal.*
(3) *An α-infinitesimal is α-finite.*
(4) *If x is α-infinitesimal and $|y| \leq x$, then y is α-infinitesimal.*
(5) *If x and y are α-finite, then so are $x+y$ and $x \times y$.*
(6) *The number ε_α is β-infinitesimal for all $\beta \prec \alpha$.*
(7) *The number $\omega_\alpha = 1/\varepsilon_\alpha$ is hypernatural and α-finite.*

2.6. Theorem. *The number ω_α is β-infinite for all $\beta \prec \alpha$.*

Proof. Immediate from items (6) and (7) of the previous axiom set. □

2.7. Theorem. *x is α-finite iff there is an α-natural n such that $|x| \leq n$.*

Proof. The statement is trivial for $x = 0$. If $x \neq 0$ is α-finite, so is $|x|$ because, assuming the opposite, $1/|x|$ would be α-infinitesimal and so would $1/x$ be by axiom 2.5.(4). By axiom 2.5.(5), the hypernatural $n = \lceil |x| \rceil < |x|+1$ is then also α-finite. Conversely, let n be α-natural and $|x| \leq n$. If $1/|x|$ were α-infinitesimal, so would $1/n$ be by axiom 2.5.(4), and this contradicts the assumption that n is α-finite. □

Thus, we see that $L^{\alpha\text{-}st}$ is just L^{st} with all α-finite constants added.

2.8. Corollary. *$x \approx_\alpha 0$ iff $|x| < 1/n$ for all α-natural $n \geq 1$.*

In the following, we assume that the function f and the formulas φ and Φ do not involve ERNA$^{\mathbb{A}}$'s minimum operator.

2.2. Consistency. In this section, we prove the consistency of ERNA$^{\mathbb{A}}$ inside PRA. As ERNA$^{\mathbb{A}}$ is a quantifier-free theory, we can use Herbrand's theorem in the same way as in [**28**], [**29**] and [**49**], for more details, see [**8**] or [**21**]. To obtain ERNA's original consistency proof from the following, omit $\approx_\alpha$ for $\alpha \neq \mathbf{0}$ from the language.

2.9. Theorem. *The theory* ERNA$^{\mathbb{A}}$ *is consistent and this consistency can be proved in* PRA.

Proof. In view of Herbrand's theorem, it suffices to show the consistency of every finite set of instantiated axioms of ERNA$^{\mathbb{A}}$. Let T be such a set. We will define a mapping val_α on T, similar to the mapping val in ERNA's consistency proof. Thus, val_α maps the terms of T to rationals and the relations of T to relations on rationals, in such a way that all axioms of T are true under val_α. Hence, T is consistent and the theorem follows.

First of all, as there are only finitely many elements of $\mathbb{A}$ in T, we interpret $(\mathbb{A}, \preceq)$ as a suitable initial segment of $(\mathbb{N}, \leq)$.

Second, like in the consistency proof of ERNA, all standard terms of T, except for min, are interpreted as their homomorphic image in the rationals: for all terms occurring in T, except min, ε_α, ω_α, we define

$$\text{val}_\alpha(f(x_1, \dots, x_k)) := f(\text{val}_\alpha(x_1), \dots, \text{val}_\alpha(x_k)) \tag{2.58}$$

and for all relations R occurring in T, except $\approx_\alpha$, we define

$$\mathrm{val}_\alpha(R(x_1,\dots,x_k)) \text{ is true} \leftrightarrow R(\mathrm{val}_\alpha(x_1),\dots,\mathrm{val}_\alpha(x_k)). \tag{2.59}$$

Third, we need to gather some technical machinery. Let D be the maximum depth of the terms in T and let $\alpha_0,\alpha_1,\alpha_2,\dots,\alpha_{N-1}$ be all numbers of $\mathbb{A}$ that occur in T, with $\alpha_0 = \mathbf{0}$. As ERNA$^{\mathbb{A}}$ has the same axiom schema for recursion as ERNA, no standard term of ERNA$^{\mathbb{A}}$ grows faster than 2_k^x, for $k \in \mathbb{N}$. Hence, by theorem 1.33, there is a $0 < B \in \mathbb{N}$ such that for every term $f(\vec{x})$ occurring in T, not involving min, we have

$$||f(\vec{x})|| \le 2_B^{||\vec{x}||}. \tag{2.60}$$

Further assume that t_D is the number of terms of depth D one can create using only function symbols occurring in T, and define $t := 3t_D + 3$.
With t and D, define the following functions:

$$f_0(x) = 2_B^x \text{ and } f_{n+1}(x) = f_n^t(x) = \underbrace{f_n(f_n(\dots(f_n(x))))}_{t\ f_n\text{'s}}. \tag{2.61}$$

Furthermore, define $a_0 := 1$ and

$$b_0^1 := f_{D+1}(a_0), c_0^1 := b_0^1, b_0^2 := f_{D+1}(c_0^1), c_0^2 := b_0^2, \dots, b_0^N := f_{D+1}\big(c_0^{N-1}\big), \tag{2.62}$$

and finally $c_0^N := b_0^N$ and $d_0 := f_{D+1}(c_0^N)$.

The numbers b_0^l allow us to interpret ε_α and ω_α:

$$\mathrm{val}_\alpha(\omega_{\alpha_1}) := b_0^1, \mathrm{val}_\alpha(\omega_{\alpha_2}) := b_0^2, \dots, \mathrm{val}_\alpha(\omega_{\alpha_{N-1}}) := b_0^{N-1} \tag{2.63}$$

and

$$\mathrm{val}_\alpha(\varepsilon_{\alpha_1}) := 1/b_0^1, \mathrm{val}_\alpha(\varepsilon_{\alpha_2}) := 1/b_0^2, \dots, \mathrm{val}_\alpha(\varepsilon_{\alpha_{N-1}}) := 1/b_0^{N-1}. \tag{2.64}$$

Hence, we have an interpretation of all terms τ of depth zero such that $|\mathrm{val}_\alpha(\tau)| \in [0,a_0] \cup [b_0^1, c_0^1] \cup \dots \cup [b_0^N, c_0^N]$. There holds, for $i = 0$ and $1 \le l \le N-1$, that

$$b_i^1 := f_{D-i+1}(a_i), b_i^{l+1} := f_{D-i+1}(c_i^l) \text{ and } d_i = f_{D-i+1}(c_i^N). \tag{2.65}$$

Then suppose that for $i \ge 0$ the numbers a_i, b_i^l, c_i^l and d_i have already been calculated and satisfy (2.65) and suppose val_α interprets all terms τ of depth i in such a way that $|\mathrm{val}_\alpha(\tau)| \in [0,a_i] \cup [b_i^1, c_i^1] \cup \dots \cup [b_i^N, c_i^N]$. We will now define a_{i+1}, b_{i+1}^l, c_{i+1}^l and d_{i+1}, which will satisfy (2.65) for $i+1$ and interpret all terms τ of depth $i+1$ in such a way that $|\mathrm{val}_\alpha(\tau)| \in [0,a_{i+1}] \cup [b_{i+1}^1, c_{i+1}^1] \cup \dots \cup [b_{i+1}^N, c_{i+1}^N]$.

In order to obtain a suitable interpretation for min, we define,

$$n_\varphi(\vec{x}) := (\mu n \le d_i)\varphi(n, \mathrm{val}_\alpha(\vec{x})). \tag{2.66}$$

Let S_{i+1} be the set of all numbers $n_\varphi(\mathrm{val}_\alpha(\vec{\tau}))$ such that $\min_\varphi(\vec{\tau})$ has depth $i+1$ and is in T.

Now observe that, due to (2.65), the intervals $[a_i, b_i^1]$, $[c_i^l, b_i^{l+1}]$ and $[c_i^N, d_i]$ can be respectively partitioned in t intervals of the form

$$[f_{D-i}^j(a_i), f_{D-i}^{j+1}(a_i)], [f_{D-i}^j(c_i^l), f_{D-i}^{j+1}(c_i^l)] \text{ and } [f_{D-i}^j(c_i^N), f_{D-i}^{j+1}(c_i^N)] \tag{2.67}$$

for $j = 0,\dots,t-1 = 3t_D+2$. Let V_{i+1} be the set of all numbers $n_\varphi(\vec{\tau})$ in S_{i+1} and all other terms $f(\vec{x})$ of T of depth at most $i+1$. Close V_{i+1} under taking the inverse and the weight, keeping in mind that $\|x\| = \|1/x\|$. Then V_{i+1} has at most $3t_D$ elements and recall that each partition in (2.67) has $3t_D + 3$ elements. Using the pigeon-hole principle, we can pick an interval, say the j_0-th one, which has empty intersection with V_{i+1}. Note that we can assume $1 \leq j_0 \leq 3t_D + 1$, because we have a surplus of three intervals. Finally we can define

$$a_{i+1} := f_{D-i}^{j_0}(a_i) \text{ and } b_{i+1}^1 := f_{D-i}^{j_0+1}(a_i) \tag{2.68}$$

The numbers b_{i+1}^l, c_{i+1}^l and d_{i+1} are defined in the same way. Hence, (2.65) holds for $i+1$. Finally, we define

$$\mathrm{val}_\alpha(\min_\varphi(\vec{x})) := (\mu n \leq c_{i+1}^N)\varphi(n, \mathrm{val}_\alpha(\vec{x})) \tag{2.69}$$

for all $\min_\varphi(\vec{\tau})$ with depth $i+1$ in T. This definition, together with (2.60), yields that val_α interprets all terms τ of depth $i+1$ in such a way that $|\mathrm{val}_\alpha(\tau)| \in [0, a_{i+1}] \cup [b_{i+1}^1, c_{i+1}^1] \cup \cdots \cup [b_{i+1}^N, c_{i+1}^N]$. Note that the latter property holds for all terms in V_{i+1}, in particular for $1/|\mathrm{val}_\alpha(\tau)|$.

After repeating this process D times, we obtain numbers a_D, b_D^l, c_D^l and d_D which allow us to interpret all terms of T. Finally, we give an interpretation to the relations $\approx_{\alpha_l}$:

$$\mathrm{val}_\alpha(\tau \approx_{\alpha_l} 0) \text{ is true} \leftrightarrow |\tau| \leq 1/b_D^{l+1}, \tag{2.70}$$

for $0 \leq l \leq N-1$. What is left is to show that under this interpretation val_α, all the axioms of T recieve the predicate true, which is done next.

Because most axioms of $\mathrm{ERNA}^{\mathbb{A}}$ hold for the rational numbers, the formulas (2.58) and (2.59) guarantee that all axioms of T have received a valid interpretation under val_α, except for axiom set 2.5 (Infinitesimals) above and ERNA's axiom set 31 (Internal minimum).

First we treat the first axiom of 'Infinitesimals'. When either is zero, there is nothing to prove. Assume $\mathrm{val}_\alpha(\sigma \approx_{\alpha_l} 0)$ and $\mathrm{val}_\alpha(\tau \approx_{\alpha_l} 0)$ are true and that $\sigma+\tau$ appears in T. By (2.70), this implies $|\mathrm{val}_\alpha(\sigma)|, |\mathrm{val}_\alpha(\tau)| \leq 1/b_D^{l+1}$ or $1/|\mathrm{val}_\alpha(\tau)|, 1/|\mathrm{val}_\alpha(\sigma)| \geq b_D^{l+1}$. But since σ and τ have depth at most $D-1$, we have $1/|\mathrm{val}_\alpha(\tau)|, 1/|\mathrm{val}_\alpha(\sigma)| \in [0, a_{D-1}] \cup [b_{D-1}^1, c_{D-1}^1] \cup \cdots \cup [b_{D-1}^N, c_{D-1}^N]$ and since there holds $a_{D-1} \leq a_D \leq b_D^{l+1} \leq b_{D-1}^{l+1}$, they must be in $\cup_{l+1 \leq k \leq N}[b_{D-1}^k, c_{D-1}^k]$. Hence, we have $1/|\mathrm{val}_\alpha(\tau)|, 1/|\mathrm{val}_\alpha(\sigma)| \geq b_{D-1}^{l+1}$ or $|\mathrm{val}_\alpha(\tau)|, |\mathrm{val}_\alpha(\sigma)| \leq 1/b_{D-1}^{l+1}$, from which $|\mathrm{val}_\alpha(\sigma+\tau)| \leq 2/b_{D-1}^{l+1} < 1/b_D^{l+1}$. This last inequality is true, since $b_D^{l+1} > 2$ and $(b_D^{l+1})^2 < b_{D-1}^{l+1}$. We have proved that $|\mathrm{val}_\alpha(\sigma+\tau)| \leq 1/b_D^{l+1}$, which is equivalent to $\mathrm{val}_\alpha(\sigma+\tau \approx_{\alpha_l} 0)$ being true. Hence, the first axiom of the set 'Infinitesimals' receives the predicate true under val_α.

The second axiom of 'Infinitesimals' is treated in the same way as the first one.

The third axiom of 'Infinitesimals' holds trivially under val, since we cannot have that $|\mathrm{val}_\alpha(\tau)| \leq 1/b_D^{l+1}$ and $1/|\mathrm{val}_\alpha(\tau)| \leq 1/b_D^{l+1}$ hold at the same time. The fact that zero is α_l-finite, is immediate by the definition of the predicate 'x is α_l-finite'.

The fourth axiom of 'Infinitesimals' holds trivially, thanks to (2.70).

The fifth axiom of 'Infinitesimals' is treated like the first and second axiom of the same set.

The sixth and seventh item of 'Infinitesimals' both follow from (2.63), (2.64) and (2.70).

Now we will treat the axioms of the schema 'Internal minimum'. First, note that the interval $[c_{i+1}^N, d_{i+1}^N]$, defined as in (2.68), has empty intersection with V_{i+1}. In particular, no term $n_\varphi(\vec{\tau})$ of T ends up in this interval. Thus, there holds, for terms $\min_\varphi$ of depth $i+1$,

$$\mathrm{val}_\alpha(\min_\varphi(\vec{\tau})) = (\mu n \leq c^N_{i+1})\varphi(n, \mathrm{val}_\alpha(\vec{\tau})) = (\mu n \leq c^N_D)\varphi(n, \mathrm{val}_\alpha(\vec{\vec{\tau}})) \tag{2.71}$$

as c^N_D is in the interval $[c^N_{i+1}, d^N_{i+1}]$. We are ready to consider items (1)-(3) of the internal minimum schema. It is clear that item (1) always holds. For item (2), assume that the antecedent holds, i.e. $\mathrm{val}_\alpha(\min_\varphi(\vec{\tau}) > 0)$ is true. By the definition of $\mathrm{val}_\alpha(\min_\varphi)$ in (2.69), the consequent $\varphi(\mathrm{val}_\alpha(\min_\varphi(\vec{\tau})), \mathrm{val}_\alpha(\vec{\tau}))$ holds too. Hence, item (2) holds. For item (3), assume that the antecedent holds, i.e. $\varphi(\mathrm{val}_\alpha(\sigma), \mathrm{val}_\alpha(\vec{\tau}))$ holds for some σ in T. This implies $\mathrm{val}_\alpha(\sigma) \leq c^N_D$ and thus there is a number $n \leq c^N_D$ such that $\varphi(n, \mathrm{val}_\alpha(\vec{\tau}))$. By (2.71), $\mathrm{val}_\alpha(\min_\varphi(\vec{\tau}))$ is the least of these and hence the formulas '$\min_\varphi(\vec{\tau}) \leq \sigma$' and '$\varphi(\min_\varphi(\vec{\tau}), \vec{\tau})$' receive a true interpretation under val_α. Thus, item (3) is also interpreted as true and we are done with this schema.

All axioms of T have received a true interpretation under val_α, hence T is consistent and, by Herbrand's theorem, ERNA$^{\mathbb{A}}$ is. Now, Herbrand's theorem is provable in $I\Sigma_1$ and this theory is Π_2-conservative over PRA (see [**8**, **21**]). As consistency can be formalized by a Π_1-formula, it follows immediately that PRA proves the consistency of ERNA$^{\mathbb{A}}$. □

Note that if we define, in (2.62), a_0 as a number larger than 1 and any c^l_0 as a number larger than b^l_0, we still obtain a valid interpretation val_α for T and the consistency proof goes through.

The choice for PRA as a 'background theory' is motivated by historical reasons. Indeed, the following corollary is immediate.

2.10. Corollary. *The consistency of* ERNA$^{\mathbb{A}}$ *can be proved in* $I\Delta_0 + \mathrm{superexp}$.

From the proof of the theorem, it is clear that the choice of $(\mathbb{A}, \preceq)$ is arbitrary, hence it is consistent with ERNA$^{\mathbb{A}}$ that $\mathbb{A}$ is dense. It is possible to make this explicit by adding the following axiom to ERNA$^{\mathbb{A}}$, for all nonzero $\alpha, \beta \in \mathbb{A}$.

$$\omega_\alpha < \omega_\beta \to \omega_\alpha < \omega_{\frac{\alpha+\beta}{2}} < \omega_\beta. \tag{2.72}$$

The notation '$\frac{\alpha+\beta}{2}$' is of course purely symbolic. This axiom receives a valid interpretation by interpreting $(\mathbb{A}, \preceq)$ as $(\mathbb{Q}^+, \leq)$.

In the following, we repeatedly need overflow and underflow. Thus, we prove it explicitly in ERNA$^{\mathbb{A}}$.

2.11. Theorem. *Let $\varphi(n)$ be an internal quantifier-free formula.*

(1) *If $\varphi(n)$ holds for every α-natural n, it holds for all hypernatural n up to some α-infinite hypernatural $\overline{n}$* ***(overflow)***.
(2) *If $\varphi(n)$ holds for every α-infinite hypernatural n, it holds for all hypernatural n from some α-natural $\underline{n}$ on* ***(underflow)***.

Both numbers $\overline{n}$ and $\underline{n}$ are given by explicit ERNA$^{\mathbb{A}}$*-formulas not involving* min.

Proof. Let ω be some α-infinite number. For the first item, define

$$\overline{n} := (\mu n \leq \omega)\neg\varphi(n+1), \tag{2.73}$$

if $(\exists n \leq \omega)\neg\varphi(n+1)$ and zero otherwise. Likewise for underflow. □

The previous theorem shows that overflow holds for all $\alpha \in \mathbb{A}$, i.e. at all levels of infinity. As no one level is given exceptional status, this seems only natural. Furthermore, one intuitively expects formulas that do not explicitly depend on a certain level to be true at all levels if they are true at one. In the following section, we investigate a general principle that transfers universal formulas to all levels of infinity.

3. ERNA$^{\mathbb{A}}$ and Transfer

3.1. ERNA$^{\mathbb{A}}$ and Stratified Transfer. In nonstandard mathematics, Transfer expresses Leibniz's principle that the 'same' laws hold for standard and nonstandard objects alike. Typically, Transfer only applies to formulas involving standard objects, excluding e.g. ERNA's cosine $\sum_{i=0}^{\omega}(-1)^i \frac{x^{2i}}{(2i)!}$. In set theoretical approaches to Nonstandard Analysis, the standard-part function 'st' applied to such an object, results in a standard object, thus solving this problem. The latter function is not available in ERNA, but 'generalized' transfer principles for objects like ERNA's cosine can be obtained (see theorems 1.70 and 1.73), at the cost of introducing '$\approx$'. Unfortunately, formulas with occurrences of the predicate '$\approx$' are always excluded from Transfer, even in the classical set-theoretical approach.

For ERNA$^{\mathbb{A}}$, we wish to obtain a transfer principle that applies to all universal formulas, possibly involving $\approx$. As an example, consider the following formula, expressing the continuity of the standard function f on $[0,1]$:

$$(\forall x, y \in [0,1])(x \approx y \to f(x) \approx f(y)). \tag{2.74}$$

Assuming (2.74), it seems only natural that if $x \approx_\alpha y$ for $\alpha \succ \mathbf{0}$, then $f(x) \approx_\alpha f(y)$. In other words, there should hold, for all $\alpha \in \mathbb{A}$,

$$(\forall x, y \in [0,1])(x \approx_\alpha y \to f(x) \approx_\alpha f(y)), \tag{2.75}$$

which is (2.74), with $\approx$ replaced with $\approx_\alpha$. Incidentally, when f is a polynomial, an easy computation shows that (2.75) indeed holds, even for polynomials in $L^{\alpha\text{-}st}$. Below, we turn this into a general principle.

2.12. Notation. Let Φ^α be a formula of $L^{\alpha\text{-}st} \cup \{\approx_\alpha\}$. Then Φ^β is Φ^α with all occurrences of $\approx_\alpha$ replaced with $\approx_\beta$.

2.13. Principle (Stratified Transfer). *Assume $\alpha \succeq \mathbf{0}$ and let Φ^α be a quantifier-free formula of $L^{\alpha\text{-}st} \cup \{\approx_\alpha\}$. There holds, for every $\beta \succ \alpha$,*

$$(\forall \vec{x})\Phi^\alpha(\vec{x}) \leftrightarrow (\forall \vec{x})\Phi^\beta(\vec{x}). \tag{2.76}$$

2.14. Principle (Weak Stratified Transfer). *Assume $\alpha \succeq \mathbf{0}$ and let $f(\vec{x}, k)$ be a function of $L^{\alpha\text{-}st}$, weakly increasing in k. For all $\beta \succ \alpha$, the following statements are equivalent*

'$f(\vec{x}, k)$ is α-infinite for all $\vec{x}$ and all α-infinite k'

and

'$f(\vec{x}, k)$ is β-infinite for all $\vec{x}$ and all β-infinite number k'.

The second transfer principle is a special case of the first. However, by the following theorem, the seemingly weaker second principle is actually equivalent to the first. We sometimes abbreviate 'for all α-infinite ω' by '$(\forall^\alpha \omega)$'.

2.15. Theorem. *In* ERNA$^{\mathbb{A}}$, *Weak Stratified Transfer is equivalent to Stratified Transfer.*

Proof. First, assume the Weak Stratified Transfer Principle and let $\Phi^{\alpha}(\vec{x})$ be as in the Stratified Transfer Principle. Replace in $\Phi^{\alpha}(\vec{x})$ all positive occurrences of $\tau_i(\vec{x}) \approx_{\alpha} 0$ with $(\forall^{\alpha\text{-}st} n_i)(|\tau_i(\vec{x})| < 1/n_i)$, where n_i is a new variable not yet appearing in $\Phi^{\alpha}(\vec{x})$. Do the same for the negative occurrences, using new variables m_i. Bringing all quantifiers in $(\forall \vec{x})\Phi^{\alpha}(\vec{x})$ to the front, we obtain

$$(\forall \vec{x})(\forall^{\alpha\text{-}st} n_1, \dots, n_l)(\exists^{\alpha\text{-}st} m_1, \dots, m_k)\Psi(\vec{x}, n_1, \dots, n_l, m_1, \dots, m_k),$$

where Ψ is quantifier-free and in $L^{\alpha\text{-}st}$. Using pairing functions, we can reduce all n_i to one variable n and reduce all m_i to one variable m. Hence, the previous formula becomes

$$(\forall \vec{x})(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)\Xi(\vec{x}, n, m),$$

where Ξ is quantifier-free and in $L^{\alpha\text{-}st}$. Fix some α-infinite number ω_1; we obtain

$$(\forall \vec{x})(\forall^{\alpha\text{-}st} n)(\exists m \leq \omega_1)\Xi(\vec{x}, n, m),$$

Applying overflow, with $\omega = \omega_1$ in (2.73), yields

$$(\forall \vec{x})(\forall n \leq \overline{n}(\vec{x}, \omega_1))(\exists m \leq \omega_1)\Xi(\vec{x}, n, m).$$

Hence, the function $\overline{n}(\vec{x}, k)$ is α-infinite for all $\vec{x}$ and α-infinite k and weakly increasing in k. By the Weak Stratified Transfer Principle, $\overline{n}(\vec{x}, k)$ is β-infinite for all $\vec{x}$ and all β-infinite k, for $\beta \succ \alpha$. Hence, for all $\vec{x}$, β-finite n and β-infinite k, we have

$$(\exists m \leq k)\Xi(\vec{x}, n, m).$$

Fix $\vec{x}_0$ and β-finite n_0. Since $(\exists m \leq k)\Xi(\vec{x}_0, n_0, m)$ holds for all β-infinite k, underflow yields $(\exists^{\beta\text{-}st} m)\Xi(\vec{x}_0, n_0, m)$. This implies

$$(\forall \vec{x})(\forall^{\beta\text{-}st} n)(\exists^{\beta\text{-}st} m)\Xi(\vec{x}, n, m).$$

Unpairing the variables n and m and bringing the quantifiers back in the formula, we obtain $(\forall \vec{x})\Phi^{\beta}(\vec{x})$. Thus, we have proved the forward implication in (2.76). In the same way, it is proved that $(\forall \vec{x})\Phi^{\beta}(\vec{x})$ implies $(\forall \vec{x})\Phi^{\alpha}(\vec{x})$, i.e., the reverse implication in (2.76), assuming the Weak Stratified Transfer Principle.

Hence, we proved that the Weak Stratified Transfer Principle implies the Stratified Transfer Principle. As the reverse implication is trivial, we are done. $\square$

By the previous theorem, it suffices to prove the consistency of ERNA$^{\mathbb{A}}$ with the Weak Stratified Transfer Principle. Instead of proving this consistency directly, we show, in the section 3.3, that Weak Stratified Transfer is equivalent to ERNA$^{\mathbb{A}}$-versions of the well-known 'regular' transfer principle for Π_3-formulas.

To conclude this section, we point to [**22**], where the importance of Stratified Transfer is discussed. Moreover, analysis developed in ERNA$^{\mathbb{A}}$ in section 4 is more elegant when Stratified Transfer is available. Also, Stratified Transfer (in some form or other) seems to be compatible with the spirit of 'strict' finitism (see [**51**]), as it merely lifts true universal formulas to higher levels.

3.2. ERNA$^{\mathbb{A}}$ and Classical Transfer. In the next section, we show that ERNA$^{\mathbb{A}}$'s version of the transfer principle for Π_3-formulas is equivalent to Stratified Transfer. To this end, we need Π_1^α-transfer, which is ERNA$^{\mathbb{A}}$'s version of universal transfer. First, we introduce a 'stratified' version of transfer for Π_1 and Σ_1-formulas for ERNA$^{\mathbb{A}}$ and show that the extended theory is consistent. The following axiom schema is ERNA$^{\mathbb{A}}$'s version of Π_1-TRANS.

2.16. Axiom schema (Stratified Π_1-transfer). *For every quantifier-free formula $\varphi(n)$ of $L^{\alpha\text{-}st}$, we have*

$$(\forall^{\alpha\text{-}st} n)\varphi(n) \to (\forall n)\varphi(n). \tag{2.77}$$

Recall that we implicitly allow *standard* parameters in φ. We denote the previous schema by Π_1^α-TRANS and its parameter-free counterpart by Π_1^α-TRANS$^-$. After the consistency proof, the reasons for the restrictions on φ will become apparent. Resolving the implication in (2.77), we see that this formula is equivalent to

$$(0 < \min{}_{\neg\varphi} \text{ is } \alpha\text{-finite}) \vee (\forall n)\varphi(n). \tag{2.78}$$

Thus, ERNA$^{\mathbb{A}} + \Pi_1^\alpha$-TRANS$^-$ is equivalent to a quantifier-free theory and we may use Herbrand's theorem to prove its consistency. To obtain the consistency proof of ERNA $+\ \Pi_1$-TRANS$^-$ from the following proof, omit $\approx_\alpha$ for $\alpha \neq \mathbf{0}$ from the language.

2.17. Theorem. *The theory* ERNA$^{\mathbb{A}} + \Pi_1^\alpha$-TRANS$^-$ *is consistent and this consistency can be proved by a finite iteration of* ERNA$^{\mathbb{A}}$*'s consistency proof.*

Proof. Despite the obvious similarities between the theories ERNA$+\Pi_1$-TRANS$^-$ and ERNA$^{\mathbb{A}}+\Pi_1^\alpha$-TRANS$^-$, the consistency proof of the former (see theorem 1.58) breaks down for the latter. The reason is that one of the explicit conditions for the consistency proof of ERNA $+\ \Pi_1$-TRANS$^-$ to work, is that φ must be in L^{st}. But in Π_1^α-TRANS$^-$, φ is in $L^{\alpha\text{-}st}$ and as such, the formula φ in (2.78) may contain the nonstandard number ω_β for $\beta \preceq \alpha$.

However, it is possible to salvage the original proof. We use Herbrand's theorem in the same way as in the consistency proof of ERNA$^{\mathbb{A}}$. Thus, let T be any finite set of instantiated axioms of ERNA$^{\mathbb{A}} + \Pi_1^\alpha$-TRANS$^-$. Leaving out the transfer axioms from T, we are left with a finite set T' of instantiated ERNA$^{\mathbb{A}}$ axioms. Let val_α be its interpretation into the rationals as in ERNA$^{\mathbb{A}}$'s consistency proof. However, nothing guarantees that the instances of Π_1^α-TRANS$^-$ in T are also interpreted as 'true' under val_α. We will adapt val_α by successively increasing the starting values defined in (2.62), if necessary. The resulting map will interpret all axioms in T as true, not just those in T'.

Let T and T' be as in the previous paragraph. Let D be the maximum depth of the terms in T. Let $\alpha_0, \ldots, \alpha_{N-1}$ be all elements of $\mathbb{A}$ in T, with $\alpha_0 = \mathbf{0}$. For notational convenience, for φ as in Π_1^α-TRANS$^-$, we shall write $\varphi(n, \vec{\tau})$ instead of $\varphi(n)$, where $\vec{\tau}$ contains all numbers occurring in φ that are not in L^{st}. Finally, let $\varphi_1(n, \vec{\tau}_1), \ldots, \varphi_M(n, \vec{\tau}_M)$ be the quantifier-free formulas whose Π_1^α-transfer axiom (2.78) occurs in T.

By (2.70), $\Omega_l := \bigcup_{l+1 \leq i \leq N} [b_D^i, c_D^i]$ is the set where val_α maps the α_l-infinite numbers. Also, $O_l := [0, a_D] \cup [b_D^1, c_D^1] \cup \cdots \cup [b_D^l, c_D^l]$ is the set where val_α maps the α_l-finite numbers. If we have, for all $i \in \{1, \ldots, M\}$ and all $l \in \{0, \ldots, N-1\}$,

$$(\exists m \in O_l)\neg\varphi_i(m, \mathrm{val}_\alpha(\vec{\tau}_i)) \vee (\forall n \in [0, a_D] \cup \Omega_0)\varphi_i(n, \mathrm{val}_\alpha(\vec{\tau}_i)), \tag{2.79}$$

we see that val_α provides a true interpretation of the whole of T, not just T', as every instance of (2.78) receives a valid interpretation, in this case. However, nothing guarantees that (2.79) holds for all $i \in \{1, \ldots, M\}$ and all $l \in \{0, \ldots, N-1\}$. Thus, assume there is an exceptional $\varphi'(n, \vec{\tau}') := \varphi_i(n, \vec{\tau}_i)$ and l_0 for which

$$(\forall m \in O_{l_0})\varphi'(m, \mathrm{val}_\alpha(\vec{\tau}')) \wedge \big(\exists n \in \big[b_D^{l_0+1}, c_D^{l_0+1}\big]\big)\neg\varphi'(n, \mathrm{val}_\alpha(\vec{\tau}')). \tag{2.80}$$

We may assume that l_0 is the least number with this property. Then (2.80) implies $(\exists n \in \Omega_{l_0})\neg\varphi'(n, \mathrm{val}(\vec{\tau}'))$, i.e. there is an '$\alpha_{l_0}$-infinite' n such that $\neg\varphi'(n, \mathrm{val}(\vec{\tau}'))$. Now choose a number $n_0 > c_D^N$ (for notational clarity, we write $a_0 = c_0^0$, for the case $l_0 = 0$) and construct a new interpretation val'_α with the same starting values as in (2.62), except for $(c_0^{l_0})' := n_0$. This val'_α continues to make the axioms in T' true and does the same with the instances in T of the axiom

$$(0 < \min\nolimits_{\neg\varphi'}(\vec{\tau}') \text{ is } \alpha_{l_0}\text{-finite}) \vee (\forall n)\varphi'(n, \vec{\tau}') \tag{2.81}$$

Indeed, if a number $n \in \Omega_{l_0}$ is such that $\neg\varphi'(n, \mathrm{val}_\alpha(\vec{\tau}'))$, the number n is interpreted by val'_α as an α_{l_0}-finite number because $n \leq c_D^N \leq (c_0^{l_0})' \leq (c_D^{l_0})'$ by our choice of $(c_0^{l_0})'$. Thus, the sentence $(\exists n \in O'_{l_0})\neg\varphi'(n, \mathrm{val}_\alpha(\vec{\tau}'))$ is true. By definition, $\vec{\tau}'$ only contains numbers ω_{α_i} for $i \leq l_0$ and (2.63) implies $\mathrm{val}_\alpha(\omega_{\alpha_i}) = b_0^i$, for $1 \leq i \leq N$. But increasing $c_0^{l_0}$ to $(c_0^{l_0})'$, as we did before, does not change the numbers $b_0^1, \ldots, b_0^{l_0}$. Hence, $\mathrm{val}_\alpha(\vec{\tau}') = \mathrm{val}'_\alpha(\vec{\tau}')$ and so $(\exists n \in O'_{l_0})\neg\varphi'(n, \mathrm{val}_\alpha(\vec{\tau}'))$ implies $(\exists n \in O'_{l_0})\neg\varphi'(n, \mathrm{val}'_\alpha(\vec{\tau}'))$. Thus, the formula $(0 < \min_{\neg\varphi'}(\vec{\tau}') \text{ is } \alpha_{l_0}\text{-finite})$ is true under val'_α and so is the whole of (2.81).

Define T'' as T' plus all instances of (2.81) occurring in T. If there is another exceptional φ_i and l_0 such that (2.80) holds, repeat this process. Note that if we increase another c_0^j for $j \geq l_0$ and construct val''_α, the latter still makes the axioms of T' true, but the axioms of T'' as well, since increasing c_0^j does not change the interpretations of the numbers ω_{α_i} for $i \leq l_0$ either. Hence, (2.81) is true under val'' for the same reason as for val'.

This process, repeated, will certainly halt: either the two lists $\{1, \ldots, M\}$ and $\{1, \ldots, N-1\}$ become exhausted or, at some earlier stage, a valid interpretation is found for T. □

The restrictions on the formulas φ admitted in (2.77) are imposed by our consistency proof. Indeed, for every α_i occurring in T, the interpretation of ω_{α_j} for $j > i$ depends on the choice of c_0^i. By our changing $c_0^{l_0}$ into $(c_0^{l_0})' > c_0^{l_0}$, formulas like (2.81) could loose their 'true' interpretation from one step to the next, if they contain such ω_j. Likewise, the changing of c_0^l can change the interpretation of $\approx_\beta$, for any $\beta \in \mathbb{A}$, and hence this predicate cannot occur in φ. The exclusion of min has, of course, a different reason: $\min_\varphi$ is only allowed in ERNA when φ does not rely on min. Finally, note that the schema Π_1^α-TRANS$^-$ is used instead of Π_1^α-TRANS in the previous theorem. This is caused by the same 'parameter issue' which affects theorem 1.58.

By contraposition, the schema Π_1^α-TRANS implies the following schema, which we denote Σ_1^α-TRANS.

2.18. Axiom schema (Stratified Σ_1-transfer). *For every quantifier-free formula $\varphi(n)$ of $L^{\alpha\text{-}st}$, we have*

$$(\exists n)\varphi(n) \to (\exists^{\alpha\text{-}st} n)\varphi(n). \tag{2.82}$$

Note that both in (2.77) and (2.82), the reverse implication is trivial. For $\varphi \in L^{\alpha\text{-}st}$, the levels $\beta \succeq \alpha$ are sometimes called the 'context' levels of φ and α is the called the 'minimial' context level, i.e. the lowest level on which all constants occurring in φ exist. In this respect, Σ_1^α-transfer expresses that true existential formulas can be pushed down to their minimal context level, which corresponds to their level of standardness.

Finally, we introduce a weaker transfer principle which only refers to certain levels of infinity, not to the totality of numbers.

2.19. PRINCIPLE. *For every quantifier-free formula φ of $L^{\alpha\text{-}st}$ and $\beta \succ \alpha$,*

$$(\forall^{\alpha\text{-}st} n)\varphi(n) \to (\forall^{\beta\text{-}st} n)\varphi(n). \tag{2.83}$$

This schema is called Π_1^β-TRANS and Σ_1^β-TRANS is defined in the same way.

3.3. Classical vs. Stratified Transfer. Here, we prove that Stratified Transfer is equivalent to a certain 'classical' transfer principle for Π_3-formulas. First, we show that a certain transfer principle for Π_3-formulas, called Π_3^α-TRANS, is sufficient to obtain Weak Stratified Transfer. We first introduce the former. Note that it is the natural extension of Π_1^α-transfer.

2.20. AXIOM SCHEMA (Stratified Π_3-transfer). *For every quantifier-free formula φ of $L^{\alpha\text{-}st}$, we have*

$$(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(\forall^{\alpha\text{-}st} k)\varphi(n,m,k) \leftrightarrow (\forall n)(\exists m)(\forall k)\varphi(n,m,k). \tag{2.84}$$

We denote this schema by Π_3^α-TRANS. We now prove that Π_3^α-transfer is sufficient to obtain Stratified Transfer.

2.21. THEOREM. *The theory* ERNA$^{\mathbb{A}}$ + Π_3^α-TRANS *proves the Weak Stratified Transfer Principle.*

PROOF. Assume $\mathbf{0} \preceq \alpha \prec \beta$ and let f be as in the Weak Stratified Transfer Principle and assume that $f(n,\vec{x})$ is α-infinite for all $\vec{x}$ and all α-infinite n. This implies that

$$(\forall \vec{x})(\forall^{\alpha\text{-}st} n)(\forall^{\alpha}\omega)(f(\omega,\vec{x}) > n).$$

Fixing $\vec{x}_0$ and α-finite n_0 and applying underflow to the formula $(\forall^\alpha\omega)(f(\omega,\vec{x}_0) > n_0)$, yields the existence of an α-finite number k_0 such that $(f(k_0,\vec{x}_0) > n_0)$. Hence, there holds

$$(\forall \vec{x})(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(f(m,\vec{x}) > n). \tag{2.85}$$

By theorem 1.52, there is a function $g(n,\vec{x})$ which calculates the least m such that $f(m,\vec{x}) > n$, for any $\vec{x}$ and α-finite n. Fix an α-infinite hypernatural ω_1 and define $h(n)$ as $\max_{\|\vec{x}\| \le \omega_1} g(n,\vec{x})$. This implies

$$(\forall^{\alpha\text{-}st} n)(\exists m \le h(n))(\forall \vec{x})(\|\vec{x}\| \le \omega_1 \to f(m,\vec{x}) > n).$$

By noting that $h(n)$ is α-finite for α-finite n, we obtain

$$(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(\forall \vec{x})(\|\vec{x}\| \le \omega_1 \to f(m,\vec{x}) > n)$$

and also

$$(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(\forall^{\alpha\text{-}st} \vec{x})(f(m,\vec{x}) > n). \tag{2.86}$$

By Π_3^α-transfer, this implies that

$$(\forall^{\beta\text{-}st} n)(\exists^{\beta\text{-}st} m)(\forall^{\beta\text{-}st} \vec{x})(f(m,\vec{x}) > n). \tag{2.87}$$

Fixing appropriate β-finite n_0 and m_0, and applying Π_1^{α}-transfer, yields

$$(\forall^{\beta\text{-}st} n)(\exists^{\beta\text{-}st} m)(\forall \vec{x})(f(m,\vec{x}) > n).$$

This formula implies that $f(k,\vec{x})$ is β-infinite for all $\vec{x}$ and all β-infinite k. The other implication in the Weak Stratified Transfer Principle is proved in the same way. $\square$

It is clear from the proof why the theorem fails for β such that $\mathbf{0} \preceq \beta \prec \alpha$. Indeed, as f may contain ω_α, we cannot apply Π_3^{α}-transfer to (2.86) for such β.

Note that (Weak) Stratified Transfer is fundamentally different from the 'regular' transfer principles, as $\approx_\alpha$ can occur in the former, but not in the latter. In this respect, it is surprising that Π_3^{α}-TRANS implies (Weak) Stratified Transfer.

However, if we consider things from the point of view of set theory, we can explain this remarkable correspondence between 'regular' and 'stratified' transfer. Internal set theory is an axiomatic approach to nonstandard mathematics (see [**39**] for details). Examples include Nelson's **IST** ([**36**]), Kanovei's **BST** ([**39**]), Péraire's **RIST** ([**40**]) and Hrbacek's **FRIST*** and **GRIST** ([**22**] and [**23**]), which inspired parts of ERNA$^{\mathbb{A}}$. These set theories are extensions of **ZFC** and most have a so called 'Reduction Algorithm'. This effective procedure applies to certain general classes of formulas and removes any predicate not in the original $\in$-language of **ZFC**. The resulting formula agrees with the original formula on standard objects. Thus, in **GRIST**, it is possible to remove the relative standardness predicate '$\sqsubseteq$' and hence transfer for formulas in the $\in$-$\sqsubseteq$-language follows from transfer for formulas in the $\in$-language. Similarly, in theorem 2.15, we show that transfer for formulas involving the relative standardness predicate $\approx_\alpha$ can be reduced to a very specific instance, involving fewer predicates $\approx_\alpha$. Later, in theorem 2.21, we prove that the remaining standardness predicates can be removed from the formula too, producing (2.86) and (2.87). Thus, we have reduced 'stratified' transfer to 'regular' transfer. In turn, it is surprising that a set-theoretical metatheorem such as the Reduction Algorithm appears in theories with strength far below **ZFC**.

To stay in the spirit of Reverse mathematics, we should find a classical transfer principle equivalent of Stratified Transfer. The schema Π_3^{α}-TRANS is not a good candidate, as it refers to the totality of all numbers, whereas Stratified Transfer does not. Indeed, nothing prohibits the existence in ERNA$^{\mathbb{A}}$ of numbers which are bigger than all ω_α for $\alpha \in \mathbb{A}$ and hence Stratified Transfer cannot say anything about these numbers, whereas Π_3^{α}-transfer can. We could add an axiom to ERNA$^{\mathbb{A}}$ which states that every x is α-finite for some $\alpha \in \mathbb{A}$, but this clashes with ERNA$^{\mathbb{A}}$'s quantifier-free nature. It it more natural to weaken Π_3^{α}-transfer to the following axiom schema called Π_3^{β}-TRANS.

2.22. Axiom schema. *For every quantifier-free formula φ of $L^{\alpha\text{-}st}$ and $\beta \succ \alpha$,*

$$(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(\forall^{\alpha\text{-}st} k)\varphi(n,m,k) \leftrightarrow (\forall^{\beta\text{-}st} n)(\exists^{\beta\text{-}st} m)(\forall^{\beta\text{-}st} k)\varphi(n,m,k). \quad (2.88)$$

The previous theorem has the following corollary concerning Π_3^{β}-transfer.

2.23. Corollary. *In* ERNA$^{\mathbb{A}}$*, the schema* Π_3^{β}-TRANS *implies Weak Stratified Transfer.*

Proof. The schema Π_3^{β}-TRANS suffices to go from (2.86) to (2.87). The corollary is then immediate from the proof of the theorem. $\square$

Here, we show that the other implication holds too, in the presence of Π_1^α-TRANS.

2.24. THEOREM. *In* ERNA$^{\mathbb{A}}$+Π_1^α-TRANS, *Weak Stratified Transfer implies* Π_3^β-TRANS.

PROOF. Let φ be a quantifier-free formula of $L^{\alpha\text{-}st}$, and let $\beta \succ \alpha$. Assume the left-hand side of (2.88) holds, i.e. $(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(\forall^{\alpha\text{-}st} k)\varphi(n,m,k)$. Fix suitable α-finite n_0 and m_0 such that $(\forall^{\alpha\text{-}st} k)\varphi(n_0,m_0,k)$. Then Π_1^α-transfer implies $(\forall k)\varphi(n_0,m_0,k)$ and there holds $(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(\forall k)\varphi(n,m,k)$. This yields

$$(\forall l)(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(\forall k \le l)\varphi(n,m,k) \tag{2.89}$$

and also

$$(\forall l)(\forall^{\alpha}\omega)(\forall^{\alpha\text{-}st} n)(\exists m \le \omega)(\forall k \le l)\varphi(n,m,k).$$

In the previous formula, fix l and α-infinite ω and apply overflow to the resulting formula. We obtain

$$(\forall l)(\forall^{\alpha}\omega)\big(\forall n \le \overline{n}(\omega,l)\big)(\exists m \le \omega)(\forall k \le l)\varphi(n,m,k), \tag{2.90}$$

and the function $\overline{n}(k,l)$ is α-infinite for all l and α-infinite k. Moreover, it does not involve min, is weakly increasing in k and part of $L^{\alpha\text{-}st}$. By the Weak Stratified Transfer Principle, $\overline{n}(k,l)$ is β-infinite for all l and β-infinite k. Thus, (2.90) implies, in particular, that

$$(\forall l)(\forall^{\beta}\omega)(\forall^{\beta\text{-}st} n)(\exists m \le \omega)(\forall k \le l)\varphi(n,m,k).$$

Now fix l_0 and β-finite n_0 to obtain

$$(\forall^{\beta}\omega)(\exists m \le \omega)(\forall k \le l_0)\varphi(n_0,m,k).$$

Underflow yields the existence of a β-finite number $\underline{N}$ such that

$$(\forall N \ge \underline{N})(\exists m \le N)(\forall k \le l_0)\varphi(n_0,m,k).$$

Note that $\underline{N}$ depends on the choice of l_0 and n_0. As $\underline{N}$ is β-finite, this implies

$$(\exists^{\beta\text{-}st} m)(\forall k \le l_0)\varphi(n_0,m,k).$$

The previous formula can be obtained for any l_0 and β-finite n_0, yielding

$$(\forall l)(\forall^{\beta\text{-}st} n)(\exists^{\beta\text{-}st} m)(\forall k \le l)\varphi(n,m,k). \tag{2.91}$$

For β-infinite l, the previous formula implies the right-hand side of (2.88). In exactly the same way, the righ-hand side of (2.88) implies the left-hand side. $\square$

2.25. COROLLARY. *In* ERNA$^{\mathbb{A}}$ + Π_1^α-TRANS, *the Stratified Transfer Principle is equivalent to* Π_3^β-TRANS.

The obvious gap we left, namely the transfer principle for Π_2-formulas, is filled now. Consider the following transfer principle, called Π_2^β-TRANS. Analogously, one defines Π_n^β-transfer.

2.26. PRINCIPLE. *For every quantifier-free formula φ of $L^{\alpha\text{-}st}$ and $\beta \succ \alpha$,*

$$(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)\varphi(n,m) \leftrightarrow (\forall^{\beta\text{-}st} n)(\exists^{\beta\text{-}st} m)\varphi(n,m). \tag{2.92}$$

By the 'very Weak Stratified Transfer Principle', we denote the weak transfer principle, limited to functions f which do not involve parameters $\vec{x}$. Thus, the former is the 'parameter-free' version of the weak transfer principle.

2.27. THEOREM. *In* ERNA$^{\mathbb{A}}$, *the very Weak Stratified Transfer Principle is equivalent to* Π_2^β-TRANS.

PROOF. Let $f(k)$ be a function of $L^{\alpha\text{-}st}$, which is weakly increasing in k. Using stratified underflow, it is easily proved, in exactly the same way as corollary 2.57, that the statement '$f(k)$ is β-infinite for all β-infinite k' is equivalent to $(\forall^{\beta\text{-}st} n)(\exists^{\beta\text{-}st} m)(f(m) > n)$, for any $\beta \in \mathbb{A}$. Thus, it is clear that the very Weak Stratified Transfer Principle is a special case of Π_2^{β}-TRANS. In exactly the same way as in theorem 2.58, Π_2^{β}-TRANS follows from the very Weak Stratified Transfer Principle. $\square$

Corollary 2.59 suggests a more fitting name for the very Weak Stratified Transfer Principle, namely 'stratified unboundedness principle'.

4. Mathematics in ERNA$^{\mathbb{A}}$

In this section, we obtain some basic theorems of analysis. We shall work in ERNA$^{\mathbb{A}}$ + Π_3^{α}-TRANS, i.e. we may use the Stratified Transfer Principle. Most theorems can be proved in ERNA$^{\mathbb{A}}$, at the cost of adding extra technical conditions. This is usually mentioned in a corollary.

For the rest of this section, we assume that $\mathbf{0} \prec \alpha \prec \beta$, that $a \ll_{\alpha} b$ are α-finite (see Notation 2.31) and that the functions f and g do not involve min.

4.0.1. *Continuity.* Here, we define the notion of continuity in ERNA$^{\mathbb{A}}$ and prove some fundamental theorems.

2.28. DEFINITION. A function f is α-continuous at a point x_0, if $x \approx_{\alpha} x_0$ implies $f(x) \approx_{\alpha} f(x_0)$. A function is α-continuous over $[a, b]$ if

$$(\forall x, y \in [a, b])(x \approx_{\alpha} y \to f(x) \approx_{\alpha} f(y)).$$

As usual, we write 'continuous' instead of '$\mathbf{0}$-continuous'. If f is α and β-continuous for $\alpha \neq \beta$, we say that f is 'α, β-continuous'

2.29. THEOREM. *If f is α-continuous over $[a, b]$ and α-finite in one point of $[a, b]$, it is α-finite for all x in $[a, b]$.*

PROOF. Let f be as in the theorem, fix α-finite k_0 and consider

$$(\forall x, y \in [a, b])(|x - y| \leq 1/N \wedge \|x, y\| \leq \omega_{\beta} \to |f(x) - f(y)| < 1/k_0). \tag{2.93}$$

As f is α-continuous, this formula holds for all α-infinite N. By theorem 1.52, (2.93) can be treated as quantifier-free and applying underflow yields that it holds for all $N \geq N_0$, where N_0 is α-finite. Then let $x_0 \in [a, b]$ be such that $f(x_0)$ is α-finite. We may assume $\|x_0\| \leq \omega_{\beta}$. Using (2.93) for $N = N_0$, it easily follows that $f(x)$ deviates at most $(N_0\lceil b - a \rceil)/k_0$ from $f(x_0)$ for $\|x\| \leq \omega_{\beta}$. As the points $x_n := a + \frac{n(b-a)}{\omega_{\beta}}$ partition the interval $[a, b]$ in α-infinitesimal subintervals, the theorem follows. $\square$

2.30. COROLLARY. *If $f \in L^{\alpha\text{-}st}$ is α-continuous over $[a, b]$, it is α-finite for all $x \in [a, b]$.*

PROOF. Let $f(x, \vec{x})$ be the function $f(x)$ from the corollary with all nonstandard numbers replaced with free variables. By theorem 1.33, there is a $k \in \mathbb{N}$ such that $\|f(x, \vec{x})\| \leq 2_k^{\|x, \vec{x}\|}$. Thus, $f(x)$ is α-finite for α-finite x. Applying the theorem finishes the proof. $\square$

By Stratified Transfer, an α-continuous function of $L^{\alpha\text{-}st}$ $\big($e.g. $\mathrm{ERNA}^{\mathbb{A}}$'s cosine $\sum_{n=0}^{\omega_\alpha}(-1)^n\frac{x^{2n}}{(2n)!}\big)$ is also β-continuous for all $\beta \succeq \alpha$. Similar statements hold for integrability and differentiability. For the sake of brevity, we mostly do not explicitly mention these properties.

4.0.2. *Differentiation.* Here, we define the notion of differentiability in $\mathrm{ERNA}^{\mathbb{A}}$ and prove some fundamental theorems. To this end, we need some notation.

2.31. Notation.

(1) A nonzero number x is '$\overline{\alpha}$-infinitesimal' or 'strict α-infinitesimal' (with respect to β) if $x \approx_\alpha 0 \wedge x \not\approx_\beta 0$. We denote this by $x \approx_{\overline{\alpha}} 0$.
(2) We write '$a \ll_\alpha b$' instead of '$a \leq b \wedge a \not\approx_\alpha b$' and '$a \lessapprox_\beta b$' instead of '$a \leq b \vee a \approx_\beta b$'.
(3) We write $\Delta_h(f)(x)$ instead of $\frac{f(x+h)-f(x)}{h}$.

We use the following notion of differentiability.

2.32. Definition.

(1) A function f is 'α-differentiable at x_0' if $\Delta_\varepsilon f(x_0) \approx_\alpha \Delta_{\varepsilon'} f(x_0)$ for all nonzero $\varepsilon, \varepsilon' \approx_\alpha 0$ and both quotients are α-finite.
(2) If f is α-differentiable at x_0 and $\varepsilon \approx_\alpha 0$, then $\Delta_\varepsilon f(x_0)$ is called 'the derivative of f at x_0' and is denoted $D_\alpha f(x_0)$.
(3) A function f is called 'α-differentiable over (a, b)' if it is α-differentiable at every point $a \ll_\alpha x \ll_\alpha b$.
(4) The concepts '$\overline{\alpha}$-differentiable' and '$\overline{\alpha}$-derivative' are defined by replacing, in the previous items, '$\varepsilon, \varepsilon' \approx_\alpha 0$' by '$\varepsilon, \varepsilon' \approx_{\overline{\alpha}} 0$'. We use the same notation for the $\overline{\alpha}$-derivative as for the α-derivative.

The choice of ε is arbitrary and hence the derivative is only defined 'up to infinitesimals'. There seems to be no good way of defining it more 'precisely', i.e. not up to infinitesimals, without the presence of a 'standard part' function 'st_α' which maps α-finite numbers to their α-standard part.

2.33. Theorem. *If a function f is α-differentiable over (a, b), it is α-continuous at all $a \ll_\alpha x \ll_\alpha b$.*

Proof. Immediate from the definition of differentiability. □

2.34. Theorem. *Let $f(x)$ and $g(x)$ be α-standard and α-differentiable over (a, b). Then $f(x)g(x)$ is α-differentiable over (a, b) and*

$$D_\alpha(fg)(x) \approx_\alpha D_\alpha f(x)\, g(x) + f(x) D_\alpha g(x) \tag{2.94}$$

for all $a \ll_\alpha x \ll_\alpha b$.

Proof. Assume f and g are α-differentiable over (a, b). Let ε be an α-infinitesimal and x such that $a \ll_\alpha x \ll_\alpha b$. Then,

$$\begin{aligned}
D_\alpha(fg)(x) &\approx_\alpha \tfrac{1}{\varepsilon}(f(x+\varepsilon)g(x+\varepsilon) - f(x)g(x)) \\
&= \tfrac{1}{\varepsilon}(f(x+\varepsilon)g(x+\varepsilon) - f(x)g(x+\varepsilon) + f(x)g(x+\varepsilon) - f(x)g(x)) \\
&= \tfrac{1}{\varepsilon}((f(x+\varepsilon) - f(x))g(x+\varepsilon) + f(x)(g(x+\varepsilon) - g(x))) \\
&= \tfrac{f(x+\varepsilon)-f(x)}{\varepsilon} g(x+\varepsilon) + f(x)\tfrac{g(x+\varepsilon)-g(x)}{\varepsilon} \\
&\approx_\alpha D_\alpha f(x) g(x+\varepsilon) + f(x) D_\alpha g(x) \approx_\alpha D_\alpha f(x) g(x) + f(x) D_\alpha g(x).
\end{aligned}$$

The final two steps follow from theorem 2.33 and corollary 2.30. Hence, $f(x)g(x)$ is α-differentiable over (a,b) and (2.94) indeed holds. $\square$

By theorem 2.29, the requirement '$f,g \in L^{\alpha\text{-}st}$' in the previous theorem, can be dropped if we additionally require fg to be α-finite in one point of (a,b). In the following theorem, there is no such requirement.

2.35. Theorem (Chain rule). *Let g be α-differentiable at a and let f be α-differentiable at $g(a)$. Then $f\circ g$ is α-differentiable at a and*

$$D_\alpha(f\circ g)(a) \approx_\alpha D_\alpha f(g(a))\, D_\alpha g(a). \tag{2.95}$$

Proof. Let f and g be as in the theorem and assume $0 \neq \varepsilon \approx_\alpha 0$. First of all, since g is α-differentiable at a, we have, that $D_\alpha g(a) \approx_\alpha \frac{g(a+\varepsilon)-g(a)}{\varepsilon}$, which implies

$$g(a+\varepsilon) = \varepsilon D_\alpha g(a) + g(a) + \varepsilon\varepsilon'$$

for some $\varepsilon' \approx_\alpha 0$. Then $\varepsilon'' = \varepsilon D_\alpha g(a) + \varepsilon\varepsilon'$ is also α-infinitesimal. If $\varepsilon'' \neq 0$, then, as f is α-differentiable at $g(a)$, we have $D_\alpha f(g(a)) \approx_\alpha \frac{f(g(a)+\varepsilon'')-f(g(a))}{\varepsilon''}$. This implies

$$f(g(a)+\varepsilon'') = \varepsilon'' D_\alpha f(g(a)) + f(g(a)) + \varepsilon''\varepsilon'''$$

for some $\varepsilon''' \approx_\alpha 0$. If $\varepsilon'' = 0$, then the previous formula holds trivially for the same ε'''. Note that $\frac{\varepsilon''\varepsilon'''}{\varepsilon} \approx_\alpha 0$. Hence, we have

$$\begin{aligned}\Delta_\varepsilon(f\circ g)(a) = \frac{f(g(a+\varepsilon))-f(g(a))}{\varepsilon} &= \frac{f(g(a)+\varepsilon'')-f(g(a))}{\varepsilon}\\ &= \frac{\varepsilon'' D_\alpha f(g(a))+\varepsilon''\varepsilon'''+f(g(a))-f(g(a))}{\varepsilon}\\ &\approx_\alpha \frac{\varepsilon''}{\varepsilon} D_\alpha f(g(a)).\end{aligned}$$

By definition, $\frac{\varepsilon''}{\varepsilon} \approx_\alpha D_\alpha g(a)$ and hence $f\circ g$ is α-differentiable at a and (2.95) holds. $\square$

It is easily verified that the theorems of this section so far still hold if we replace 'α-differentiable' with '$\overline{\alpha}$-differentiable'.

As in section 3.1, we expect ERNA$^{\mathbb{A}}$'s derivative to be continuous.

2.36. Theorem. *If f is α-differentiable over (a,b), then $D_\alpha f(x)$ is α-continuous over (a,b).*

Proof. Choose points $x \approx_\alpha y$ such that $a \ll_\alpha x < y \ll_\alpha b$. If $|x-y| = \varepsilon \approx_\alpha 0$, then

$$\Delta_\varepsilon f(x) = \frac{f(x+\varepsilon)-f(x)}{\varepsilon} = \frac{f(y)-f(y-\varepsilon)}{\varepsilon} = \frac{f(y-\varepsilon)-f(y)}{-\varepsilon} = \Delta_{-\varepsilon} f(y) \approx_\alpha \Delta_\varepsilon f(y),$$

and thus $D_\alpha f(x) \approx_\alpha \Delta_\varepsilon f(x) \approx_\alpha \Delta_\varepsilon f(y) \approx_\alpha D_\alpha f(y)$. $\square$

The theorem generalizes to $\overline{\alpha}$-differentiable functions, in an elegant way.

2.37. Corollary. *If f is $\overline{\alpha}$-differentiable over (a,b), then $D_\alpha f(x)$ is α-continuous over (a,b).*

Proof. Choose $x \approx_\alpha y$ such that $a \ll_\alpha x < y \ll_\alpha b$. First, suppose $|x-y| = \varepsilon \approx_{\overline{\alpha}} 0$. The same proof as in the theorem yields this case. Now suppose we do not have $|x-y| = \varepsilon \approx_{\overline{\alpha}} 0$ and define $z = y + 2\varepsilon'$ with $\varepsilon' \approx_{\overline{\alpha}} 0$ nonzero. Then $|z-x| = \varepsilon'' \approx_{\overline{\alpha}} 0$ and $|z-y| = \varepsilon''' \approx_{\overline{\alpha}} 0$ and by the previous case, we have $\Delta_{\varepsilon''} f(x) \approx_\alpha \Delta_{\varepsilon''}(f)(z)$ and $\Delta_{\varepsilon'''} f(z) \approx_\alpha \Delta_{\varepsilon'''}(f)(y)$. As f is $\overline{\alpha}$-differentiable, we have $\Delta_{\varepsilon''}(f)(z) \approx \Delta_{\varepsilon'''}(f)(z)$ and hence $D_\alpha f(x) \approx_\alpha \Delta_{\varepsilon''} f(x) \approx_\alpha \Delta_{\varepsilon'''} f(y) \approx_\alpha D_\alpha f(y)$. $\square$

4.0.3. *Integration.* Here, we define the notion of Riemann integral in ERNA$^{\mathbb{A}}$ and prove some fundamental theorems.

In classical analysis, the Riemann-integral is defined as the limit of Riemann sums over ever finer partitions. In ERNA$^{\mathbb{A}}$, we adopt the following definition for the concept 'partition'.

2.38. Definition. A partition π of $[a, b]$ is a vector $(x_1, \dots, x_n, t_1, \dots t_{n-1})$ such that $x_i \leq t_i \leq x_{i+1}$ for all $1 \leq i \leq n-1$ and $a = x_1$ and $b = x_n$. The number $\delta_\alpha = \max_{2 \leq i \leq n}(x_i - x_{i-1})$ is called the 'mesh' of the partition π.

A partition π is called 'α-fine' if $\delta_\pi \approx_\alpha 0$. Assume that ω is α-infinite and that $a \ll_\alpha b$. Let n_0 be the least n such that $\frac{n}{\omega} > a$ and let n_1 be the least n such that $\frac{n}{\omega} > b$. Define $a_\omega := \frac{n_0}{\omega}$ and $b_\omega := \frac{n_1 - 1}{\omega}$. Like the derivative, the Riemann integral can only be defined 'up to infinitesimals'. Hence, for α-Riemann integrable functions, it does not matter whether we use the interval $[a, b]$ or the interval $[a_\omega, b_\omega]$ in its definition.

2.39. Definition (Riemann Integration). Let f be a function defined on $[a, b]$.

(1) Given a partition $(x_1, \dots, x_n, t_1, \dots, t_{n-1})$ of $[a, b]$, the Riemann sum corresponding to f is defined as $\sum_{i=2}^{n} f(t_{i-1})(x_i - x_{i-1})$.
(2) The function f is α-Riemann integrable on $[a, b]$, if for all partitions of $[a, b]$ with mesh $\approx_\alpha 0$, the Riemann sums are α-finite and α-infinitely close.
(3) If f is α-Riemann integrable on $[a, b]$ and π is an α-fine partition of $[a, b]$, then $\int_a^b f(x)\, d_\pi(x, \alpha)$, the integral of f over $[a, b]$, is the Riemann sum corresponding to f and π.

Note that the integral $\int_a^b f\, d(x, \alpha)$ is only defined up to α-infinitesimals, as expected. We treat it in the same way as in Chapter I.

2.40. Theorem. *A function f which is α-continuous and α-finite over $[a, b]$, is α-Riemann integrable over $[a, b]$.*

Proof. The proof for $\alpha = \mathbf{0}$ in Chapter I is easily adapted to $\alpha \succ \mathbf{0}$. $\square$

2.41. Theorem. *Let f be α-continuous and α-finite over $[a, b]$ and assume $a \ll_\alpha c \ll_\alpha b$. We have*

$$\int_a^b f(x)\, d(x, \alpha) \approx_\alpha \int_a^c f(x)\, d(x, \alpha) + \int_c^b f(x)\, d(x, \alpha).$$

Proof. Immediate from the previous theorem and definition 2.39. $\square$

2.42. Theorem. *Let c be an α-finite positive constant such that $c \not\approx_\alpha 0$ and let f be α-continuous and α-finite over $[a, b + c]$. We have*

$$\int_a^b f(x + c)\, d(x, \alpha) \approx_\alpha \int_{a+c}^{b+c} f(x)\, d(x, \alpha).$$

Proof. Immediate from theorem 2.40 and the definition of the Riemann integral. $\square$

2.43. THEOREM (First fundamental theorem). *Let $f \in L^{\alpha\text{-}st}$ be α-continuous on $[a,b]$ and let $F(x)$ be $\int_a^x f(t)d(t,\beta)$. Then $F(x)$ is $\overline{\alpha}$-differentiable over (a,b) and the equation $D_\alpha F(x) \approx_\alpha f(x)$ holds for all $a \ll_\alpha x \ll_\alpha b$.*

PROOF. Fix $\varepsilon \approx_{\overline{\alpha}} 0$ and x such that $a \ll_\alpha x \ll_\alpha b$. We have

$$\frac{F(x+\varepsilon)-F(x)}{\varepsilon} = \frac{1}{\varepsilon}\left(\int_a^{x+\varepsilon} f(t)d(t,\beta) - \int_a^x f(t)d(t,\beta)\right) \approx_\beta \frac{1}{\varepsilon}\int_x^{x+\varepsilon} f(t)d(t,\beta), \quad (2.96)$$

as ε is not β-infinitesimal. Let ω_1 be β-infinite and define $x_i = x + \frac{i\varepsilon}{\omega_1}$. Let $f(y_1)$ and $f(y_2)$ be the least and the largest $f(x_i)$ for $i \leq \omega_1$. As f is α,β-continuous, $m := f(y_1)$ and $M := f(y_2)$ are such that $m \lessapprox_\beta f(y) \lessapprox_\beta M$ for $y \in [x, x+\varepsilon]$ and $m \approx_\alpha M \approx_\alpha f(x)$. This implies

$$\varepsilon m \lessapprox_\beta \int_x^{x+\varepsilon} f(t)d(t,\beta) \lessapprox_\beta \varepsilon M,$$

and hence

$$m \lessapprox_\beta \frac{1}{\varepsilon}\int_x^{x+\varepsilon} f(t)d(t,\beta) \lessapprox_\beta M,$$

as ε is not β-infinitesimal. Thus,

$$m \approx_\alpha \frac{1}{\varepsilon}\int_x^{x+\varepsilon} f(t)d(t,\beta) \approx_\alpha M \approx_\alpha f(x).$$

By (2.96), F is $\overline{\alpha}$-differentiable and the theorem follows. ☐

2.44. COROLLARY. *The condition '$f \in L^{\alpha\text{-}st}$' in the theorem can be dropped if we require f to be α,β-continuous over $[a,b]$ and α-finite in one point of $[a,b]$.*

PROOF. It is an easy verification that the proof of the theorem still goes through with these conditions. ☐

2.45. EXAMPLE. Define $\varepsilon = \varepsilon_\alpha^4$. The function $d(x) = \frac{\varepsilon}{\varepsilon^2+x^2}$ is α,β-continuous for α-finite x and at most $1/\varepsilon_\alpha^4$. The function $\arctan x := \int_0^x \frac{d(t,\beta)}{1+t^2}$ is $\overline{\alpha}$-differentiable in all α-finite x and we have $D_\alpha\big(\arctan(x/\varepsilon)\big) \approx_\alpha \frac{\varepsilon}{\varepsilon^2+x^2}$ for all α-finite x.

2.46. THEOREM (Second fundamental theorem). *Let $f \in L^{\alpha\text{-}st}$ be $\overline{\alpha}$-differentiable over (a,b) and such that $D_\alpha f$ is β-continuous over (a,b). For $a \ll_\alpha c \ll_\alpha d \ll_\alpha b$, we have $\int_c^d D_\alpha f(x)\,d(x,\beta) \approx_\alpha f(d) - f(c)$.*

PROOF. Let c, d be as stated and let ε be strict α-infinitesimal. Note that $d-c$ is α-finite. We have

$$\begin{aligned}
\int_c^d D_\alpha f(x)\,d(x,\beta) &\approx_\alpha \int_c^d \frac{f(x+\varepsilon)-f(x)}{\varepsilon}\,d(x,\beta) \\
&\approx_\beta \frac{1}{\varepsilon}\left(\int_c^d f(x+\varepsilon)\,d(x,\beta) - \int_c^d f(x)\,d(x,\beta)\right) \\
&\approx_\beta \frac{1}{\varepsilon}\left(\int_{c+\varepsilon}^{d+\varepsilon} f(x)\,d(x,\beta) - \int_c^d f(x)\,d(x,\beta)\right) \\
&\approx_\beta \frac{1}{\varepsilon}\left(\int_d^{d+\varepsilon} f(x)\,d(x,\beta) - \int_c^{c+\varepsilon} f(x)\,d(x,\beta)\right).
\end{aligned}$$

As in the proof of the first fundamental theorem, we have $\int_c^{c+\varepsilon} f(x)\,d(x,\beta) \approx_\alpha f(c)$ and $\int_d^{d+\varepsilon} f(x)\,d(x,\beta) \approx_\alpha f(d)$ and we are done. ☐

2.47. COROLLARY (Integration by parts). *Let $f, g \in L^{\alpha\text{-}st}$ be $\overline{\alpha}$-differentiable over (a,b) and let $D_\alpha f$ and $D_\alpha g$ be β-continuous over (a,b). For $a \ll_\alpha c \ll_\alpha d \ll_\alpha b$,*

$$\int_c^d f(x) D_\alpha g(x)\, d(x,\beta) \approx_\alpha \big[f(x)g(x)\big]_c^d - \int_c^d D_\alpha f(x) g(x)\, d(x,\beta).$$

PROOF. Immediate from the second fundamental theorem and theorem 2.34. ☐

By theorem 2.93, we can drop the requirement '$f, g \in L^{\alpha\text{-}st}$' if we additionally require fg to be α-finite in one point of (a,b).

For simulating the Dirac Delta distribution, we need to introduce an extra level γ such that $\mathbf{0} \prec \gamma \prec \alpha$. We also need the following properties of $\arctan x$, defined in example 2.45.

2.48. THEOREM. *Define the (finite) constant π as $4\arctan(1)$.*

(1) *For all α-finite x,* $\arctan(\pm|x|) + \arctan\big(\pm\frac{1}{|x|}\big) \approx_\alpha \pm\pi/2$.
(2) *We have* $\arctan(\pm\omega_\alpha^3) \approx_\gamma \pm\pi/2$.

PROOF. The first item follows by calculating the $\overline{\alpha}$-derivative of $\arctan x + \arctan 1/x$ using the chain rule and noting that the result is α-infinitely close to zero. Thus, there is a constant C such that $\arctan x + \arctan 1/x \approx_\alpha C$, for all α-finite positive x. Substituting $x = 1$ yields $C = \pi/2$. The case $x < 0$ is treated in the same way. The second item follows from the previous item and the fact that $\arctan x$ is continuous at zero. ☐

2.49. DEFINITION. A function $f \in L^{\gamma\text{-}st}$ is said to have a 'compact support' if it is zero outside some interval $[a,b]$ with γ-finite a, b.

2.50. THEOREM. *Let $f \in L^{\gamma\text{-}st}$ be a γ-differentiable function with compact support such that $D_\alpha f(x)$ is β-continuous for $x \approx_\gamma 0$. We have*

$$\frac{1}{\pi}\int_{-\omega_\alpha}^{\omega_\alpha} d(x) f(x) d(x,\beta) \approx_\gamma f(0).$$

PROOF. Assume that $f(x)$ is zero outside $[a,b]$, with γ-finite a,b. First, we prove that $\int_{\varepsilon_\alpha}^b f(x)d(x)\, d(x,\beta) \approx_\gamma 0$. As $|x| \geq \varepsilon_\alpha$ implies $x^2 \geq \varepsilon_\alpha^2$ we have $d(x) = \frac{\varepsilon}{\varepsilon^2+x^2} \leq \frac{\varepsilon}{x^2} \leq \frac{\varepsilon}{{\varepsilon_\alpha}^2} = \varepsilon_\alpha^2 < \varepsilon_\alpha$. Hence, the integral $\int_{\varepsilon_\alpha}^b |d(x)|\,|f(x)|\, d(x,\beta)$ is at most $\varepsilon_\alpha \int_{\varepsilon_\alpha}^b |f(x)|\, d(x,\beta)$. As f is γ-finite and γ-continuous on $[a,b]$, we have $\int_{\varepsilon_\alpha}^b f(x)d(x)\, d(x,\beta) \approx_\gamma 0$. In the same way, we have $\int_a^{\varepsilon_\alpha} f(x)d(x)\, d(x,\beta) \approx_\gamma 0$ and $\int_{-\varepsilon_\alpha}^{\varepsilon_\alpha} \arctan(x/\varepsilon)) D_\alpha f(x)\, d(x,\beta) \approx_\gamma 0$. For the final integral, note that $D_\alpha f(x)$ is β-continuous and γ-finite, by assumption and that $\arctan x$ is finitely bounded for α-finite x. Hence, we have

$$\int_{-\omega_\alpha}^{\omega_\alpha} d(x) f(x) d(x,\beta) \approx_\beta \int_a^b d(x) f(x)\, d(x,\beta) \approx_\gamma \int_{-\varepsilon_\alpha}^{\varepsilon_\alpha} d(x) f(x)\, d(x,\beta).$$

If $0 \notin [a,b]$, then $f(0) = 0$ and the theorem follows. Otherwise, by example 2.45, the function $d(x)$ is α-infinitely close to $D_\alpha \arctan(x/\varepsilon)$, yielding

$$\int_{-\varepsilon_\alpha}^{\varepsilon_\alpha} d(x)f(x)\,d(x,\beta) \approx_\alpha \int_{-\varepsilon_\alpha}^{\varepsilon_\alpha} D_\alpha(\arctan(x/\varepsilon))\,f(x)\,d(x,\beta).$$

The product $\arctan(x/\varepsilon)f(x)$ satisfies all conditions for integration by parts, implying

$$\begin{aligned}
&\int_{-\varepsilon_\alpha}^{\varepsilon_\alpha} D_\alpha(\arctan(x/\varepsilon))\,f(x)\,d(x,\beta)\\
&\approx_\alpha \big[\arctan(x/\varepsilon)\;\; f(x)\big]_{-\varepsilon_\alpha}^{\varepsilon_\alpha} - \textstyle\int_{-\varepsilon_\alpha}^{\varepsilon_\alpha} \arctan(x/\varepsilon_\alpha) D_\alpha f(x)\,d(x,\beta)\\
&\approx_\gamma [\arctan(x/\varepsilon)\;\; f(x)]_{-\varepsilon_\alpha}^{\varepsilon_\alpha}\\
&= \big(\arctan(\varepsilon_\alpha/\varepsilon)\,f(\varepsilon_\alpha) - \arctan(-\varepsilon_\alpha/\varepsilon)\,f(-\varepsilon_\alpha)\big)\\
&\approx_\gamma \big(\arctan(\omega_\alpha^3)\;\, f(0) - \arctan(-\omega_\alpha^3)\;\, f(0)\big) \approx_\gamma \pi f(0).
\end{aligned}$$

□

The function $d(x)$ has the typical 'Dirac Delta' shape: 'infinite at zero and zero everywhere else' and many functions like $d(x)$ exist. Also, if we define $H(x) = \frac{1}{\pi}\arctan(x/\varepsilon) + \frac{1}{2}$, we have $D_\alpha H(x) \approx_\alpha d(x)$ and $H(x)$ only differs from the 'usual' Heaviside function by an infinitesimal. In the same way as in the previous theorem, it is possible to prove statements like

$$\int_{-\omega_\alpha}^{\omega_\alpha} D_\xi d(x) f(x) d(x,\beta) \approx_\gamma -\int_{-\omega_\alpha}^{\omega_\alpha} d(x) D_\xi f(x) d(x,\beta) \approx_\gamma -\pi D_\xi f(0).$$

in ERNA$^{\mathbb{A}}$, for $\alpha \prec \xi \prec \beta$. We have introduced the function $\arctan x$, because we needed its properties in theorem 2.50. The rest of the basic functions of analysis are easily defined and their well-known properties are almost immediate, thanks to Stratified Transfer.

In this section, we have shown that analysis can be developed inside ERNA$^{\mathbb{A}}$ and its extensions in a concise and elegant way. We did not attempt to give an exhaustive treatment and have deliberately omitted large parts of analysis like e.g. higher order derivatives. It is interesting, however, to briefly consider the latter. In [**22**], Hrbacek argues that stratified analysis yields a more elegant way of defining higher order derivatives than regular Nonstandard Analysis. In this way, a function $D_\alpha f(x)$ is differentiable, if it is β-differentiable for $\beta \succ \alpha$ and $f''(x)$ is defined as $D_\beta D_\alpha f(x)$. Thus, to manipulate an object such as $D_\alpha f(x)$, which is not part of $L^{\alpha\text{-}st}$, we need to go to a higher level β, where $D_\alpha f(x)$ is standard. The same principle is at the heart of most theorems in this section. This principle is the essence of stratified analysis, and occurs in all of mathematics: to study a set of objects, we extend it and gain new insights (e.g. real versus complex analysis). Thanks to Stratified Transfer, all levels have the same standard properties and thus, the extension to a higher level is always uniform.

In conclusion, we note that the Reverse Mathematics of Chapter I generalizes almost trivially to ERNA$^{\mathbb{A}}$. In particular, we can formulate an elegant version of the

Bolzano-Weierstraß theorem. Let $\mathbb{N}^\alpha$ be the set of α-finite numbers and assume $\alpha \prec \gamma \prec \beta$.

2.51. Theorem (Stratified Bolzano-Weierstraß). *For every $\tau(n) \in L^{\alpha\text{-}st}$, α-finitely bounded on $\mathbb{N}^\alpha$, there is a function $\sigma : \mathbb{N}^\alpha \to \mathbb{N}^\alpha$ such that $\tau(\sigma(n, \omega_\beta))$ converges to any $\tau(\sigma(m, \omega_\beta))$ with m α-infinite and γ-finite. For such m, m', we have $\tau(\sigma(m, \omega_\beta)) \approx_\alpha \tau(\sigma(m', \omega_\beta))$.*

4.0.4. *A formal framework for Physics.* We have introduced ERNA$^{\mathbb{A}}$ and proved its consistency inside PRA. We subsequently obtained several results of analysis using the elegant framework of stratified analysis. Thus, ERNA$^{\mathbb{A}}$ is a good formal framework for doing finitistic analysis in a quantifier-free way, akin to the way mathematics is done in Physics. As it turns out, Stratified Transfer gives us an even better framework. How this works is discussed in this paragraph.

It seems only fair to say that physicists employ a lower standard of mathematical rigour than mathematicians (see [**12**]). In this way, limits are usually pushed inside or outside integrals without a second thought. Moreover, a widely held 'rule of thumb' is that if, after performing a mathematically dubious manipulation, the result still makes physical and (to a lesser extent) mathematical sense, the manipulation was probably sound. As it turns out, stratified Nonstandard Analysis is a suitable formal framework for this sort of 'justification a posteriori'. We illustrate this with an example.

2.52. Example. Let f_i, a and b be standard objects. According to the previously mentioned 'rule of thumb', the following manipulation

$$\int_a^b \sum_{i=0}^{\infty} f_i(x,y)\,dx = \sum_{i=0}^{\infty} \int_a^b f_i(x,y)\,dx =: \sum_{i=0}^{\infty} g_i(y) =: g(y)$$

is considered valid in Physics as long as the function $g(y)$ is physically and/or mathematically meaningful. In stratified analysis, assuming $\mathbf{0} \prec \alpha \prec \beta$, the previous becomes

$$\int_a^b \sum_{i=0}^{\omega_\alpha} f_i(x,y)\,d(x,\beta) \approx \sum_{i=0}^{\omega_\alpha} \int_a^b f_i(x,y)\,d(x,\beta) =: \sum_{i=0}^{\omega_\alpha} h_i(y) =: h(y).$$

The first step follows from Stratified Transfer. Indeed, as a finite summation can be pushed through a Riemann integral, a β-finite summation can be pushed through a β-Riemann integral. Thus, we can always obtain $h(y)$ and if it is finite (the very least for it to be physically meaningful), there holds $h(y) \approx g(y)$, thus justifying our 'rule of thumb'.

5. ERNA$^{\mathbb{A}}$ versus ERNA

In the previous section, we obtained results for ERNA$^{\mathbb{A}}$ concerned with both logic and analysis. Here, we show that these results also yield new insights and results for ERNA. The latter would not have been obtained without our research into ERNA$^{\mathbb{A}}$. Thus, we repeat our credo:

> *The stratified nonstandard framework is a refinement of the classical one, not a departure from it.*

5.1. More Reverse Mathematics in ERNA. In this section, we discuss equivalent formulations of ERNA's transfer principle for Π_2 and Π_3-formulas. As it turns out, the result for Π_2-transfer yields a consistency proof of ERNA + Π_2-TRANS, see theorem 2.71 and corollary 2.72. Moreover, corollary 2.59 turns out to be a key element in the proof of theorem 1.161. The latter is a first step towards a 'copy up to infinitesimals' of ACA_0.

As usual, we assume that the formulas φ and Φ and the functions f and f_i do not involve ERNA's minimum operator.

5.1.1. *The Π_2-transfer principle.* Here, we show that ERNA's Π_2-transfer principle can be reduced to the following special case.

2.53. Principle. *For every weakly increasing function f in L^{st}, we have*

$$(\forall^{st} n)(\exists^{st} m)(f(m) > n) \to (\forall n)(\exists m)(f(m) > n). \tag{2.97}$$

This schema expresses that if $f(n)$ is unbounded on $\mathbb{N}$, it is unbounded everywhere. Therefore, we refer to it as the 'unboundedness principle'.

We recall ERNA's Π_2-transfer principle, which was first introduced in [**29**].

2.54. Principle (Π_2-TRANS). *For every quantifier-free formula φ in L^{st}, we have*

$$(\forall^{st} n)(\exists^{st} m)\varphi(n,m) \leftrightarrow (\forall n)(\exists m)\varphi(n,m). \tag{2.98}$$

Although the unboundedness principle seems weaker than Π_2-TRANS, they are in fact equivalent, see theorem 2.58 and corollary 2.59.

We repeatedly need two technical corollaries concerning overflow and underflow in ERNA. First, we recall the latter and then prove the corollaries.

2.55. Theorem. *Let φ be an internal quantifier-free formula.*

(1) *If $\varphi(n)$ holds for every $n \in \mathbb{N}$, it holds for all n up to some infinite $\overline{n}$* ***(overflow)***.
(2) *If $\varphi(n)$ holds for every infinite n, it holds for all n from some $\underline{n} \in \mathbb{N}$ on* ***(underflow)***.

Both numbers $\overline{n}$ and $\underline{n}$ are given by explicit ERNA*-formulas not involving* min.

Proof. For the first item, define

$$\overline{n}(k) := (\mu n \le k)\neg\varphi(n+1), \tag{2.99}$$

if $(\exists n \le k)\neg\varphi(n+1)$ and k otherwise. Then $\overline{n}(\omega)$, for infinite ω, satisfies the required properties. It is easy to prove that this term is available in ERNA. Likewise for underflow. □

Note that if φ has additional free variables, $\overline{n}$ depends on those. We are ready to prove the technical corollaries.

2.56. Corollary. *Let $\varphi \in L^{st}$ be a quantifier-free formula and let $\overline{n}(\omega)$ be obtained by applying overflow to $(\forall^{st} n)(\exists m \le \omega)\varphi(n,m)$. Then $\overline{n}(k)$ is a standard unary function.*

Proof. Formula (2.99) implies that $\overline{n}(k) = (\mu n \le k)(\forall m \le k)\neg\varphi(n+1,m)$, if $(\exists n \le k)(\forall m \le k)\neg\varphi(n+1,m)$, and zero otherwise. As there are explicit formulas for ERNA's bounded minimum and definition by cases, the theorem is immediate. □

2.57. Corollary. *Let f be weakly increasing. In* ERNA*, '$(\forall^{st}n)(\exists^{st}m)(f(m) > n)$' is equivalent to '$f(k)$ is infinite for infinite k'.*

Proof. Let f be weakly increasing. Assume that $(\forall^{st}n)(\exists^{st}m)(f(m) > n)$. Then, clearly $f(k)$ is infinite for infinite k. Now assume $f(k)$ is infinite for infinite k. This implies $(\forall^{st}n)(\forall\omega)(f(\omega) > n)$. Now fix $n_0 \in \mathbb{N}$ and apply underflow to the resulting formula. Thus, there is $\underline{m} \in \mathbb{N}$, which depends on n_0, such that $(\forall m \geq \underline{m})(f(m) > n_0)$. This yields $(\forall^{st}n)(\exists^{st}m)(f(m) > n)$. $\square$

We are ready to prove the main theorem of this paragraph.

2.58. Theorem. *In* ERNA*, the unboundedness principle implies* Π_2-TRANS*.*

Proof. We first prove the forward implication of (2.98). Assume the left-hand side of this formula holds, i.e. we have $(\forall^{st}n)(\exists^{st}m)\varphi(n,m)$. This implies, for infinite ω, that $(\forall^{st}n)(\exists m \leq \omega)\varphi(n,m)$. Applying overflow to the latter yields $(\forall\omega)(\forall n \leq \overline{n}(\omega))(\exists m \leq \omega)\varphi(n,m)$, and the function $\overline{n}(k)$ is infinite for infinite k and weakly increasing. By corollary 2.57, there holds $(\forall^{st}k)(\exists^{st}k')(\overline{n}(k') > k)$. By corollary 2.56, the function $\overline{n}(k)$ is standard, and the unboundedness principle yields $(\forall k)(\exists k')(\overline{n}(k') > k)$. By definition, the function $\overline{n}(k)$ is the largest number $n' \leq k$ such that there holds $(\forall n \leq n')(\exists m \leq k)\varphi(n,m)$. This implies $(\forall n)(\exists m)\varphi(n,m)$ and we are done.

Second, we prove the reverse implication of (2.98). This implication follows immediately if Π_1-transfer is available. Indeed, assume that the right-hand side of (2.98) holds and fix $n_0 \in \mathbb{N}$. Apply Σ_1-transfer to $(\exists m)\varphi(n_0,m)$ to obtain $(\exists^{st}m)\varphi(n_0,m)$. This yields $(\forall^{st}n)(\exists^{st}m)\varphi(n,m)$, i.e. the left-handside of (2.98). We now show that the forward implication of (2.98), which was proved above, implies Π_1-TRANS. Let $\psi \in L^{st}$ be a quantifier-free formula not involving min, and assume $(\forall^{st}n)\psi(n)$. This implies $(\forall^{st}n)(\exists^{st}m)(m > n \wedge \psi(n))$ and the forward implication of (2.98) yields $(\forall n)(\exists m)(m > n \wedge \psi(n))$. Thus, we have $(\forall n)\psi(n)$ and Π_1-TRANS follows. $\square$

2.59. Corollary. *In* ERNA*, the unboundedness principle is equivalent to* Π_2-TRANS*.*

Note that the proof of the theorem implies that we may assume that all functions occurring in the unboundedness principle are non-negative.

The equivalence proved in the previous theorem serves as a 'jumping board' to the equivalence we shall obtain for Π_3-transfer in the next paragraph.

5.1.2. *The Π_3-transfer principle.* Here, we show that transfer for Π_3-formulas is equivalent to the following 'uniform' version of the unboundedness principle. For this, we require a slightly stronger version of ERNA, defined below.

2.60. Principle (Uniform Unboundedness). *For every $f(m,\vec{x}) \in L^{st}$, weakly increasing in m,*

$$(\forall\vec{x})(\forall^{st}n)(\exists^{st}m)(f(m,\vec{x}) > n) \to (\forall n)(\exists m)(\forall\vec{x})(f(m,\vec{x}) > n). \tag{2.100}$$

The following schema is ERNA's version of transfer for Π_3-formulas.

2.61. Principle (Π_3-TRANS). *For every quantifier-free formula φ in L^{st},*

$$(\forall^{st}n)(\exists^{st}m)(\forall^{st}k)\varphi(n,m,k) \leftrightarrow (\forall n)(\exists m)(\forall k)\varphi(n,m,k). \tag{2.101}$$

For brevity, we write '*uu*-principle' or 'UUP' for the uniform unboundedness principle. It expresses that a function $f(n,\vec{x})$, unbounded over $\mathbb{N}$ for all $\vec{x}$, must be unbounded everywhere, *independent* of $\vec{x}$. The independence of $\vec{x}$ is crucial, as Π_2-transfer immediately implies that

$$(\forall \vec{x})(\forall^{st} n)(\exists^{st} m)(f(m,\vec{x}) > n) \to (\forall \vec{x})(\forall n)(\exists m)(f(m,\vec{x}) > n). \tag{2.102}$$

Thus, (2.102) is implied by the unboundedness principle, which is equivalent to Π_2-transfer. In contrast, the *uu*-principle is equivalent to Π_3-transfer, by theorem 2.62, if we slightly increase the strength of ERNA.

Let BERNA be the theory ERNA plus the replacement schema for internal Δ_0-formulas. The theory BERNA is not significantly stronger than ERNA, because $B\Sigma_1$ is Π_2-conservative over $I\Delta_0$ and $I\Sigma_1$ proves the consistency of $B\Sigma_1$ ([**8**]). Moreover, it is well-known that WKL_0^* has the same first-order strength as BERNA (see [**1**, **46**] for details). Thus, BERNA is natural from the point of view of Reverse Mathematics.

2.62. Theorem. *In* BERNA*,* UUP *implies* Π_3-TRANS.

Proof. First, note that UUP implies the unboundedness principle, and the latter is equivalent to Π_2-TRANS, by corollary 2.59. Thus, we may use Π_2-transfer (and Π_1-transfer) in this proof.

We now prove the forward implication of (2.101). Assume the left-hand side of this formula holds, i.e. we have $(\forall^{st} n)(\exists^{st} m)(\forall^{st} k)\varphi(n,m,k)$. Fix suitable $n_0, m_0 \in \mathbb{N}$ such that $(\forall^{st} k)\varphi(n_0,m_0,k)$. Then Π_1-transfer implies $(\forall k)\varphi(n_0,m_0,k)$ and there holds $(\forall^{st} n)(\exists^{st} m)(\forall k)\varphi(n,m,k)$. This yields $(\forall l)(\forall^{st} n)(\exists^{st} m)(\forall k \le l)\varphi(n,m,k)$ and also $(\forall l)(\forall \omega)(\forall^{st} n)(\exists m \le \omega)(\forall k \le l)\varphi(n,m,k)$. Fix l and infinite ω, and apply overflow to the resulting formula. We obtain

$$(\forall l)(\forall \omega)\big(\forall n \le \overline{n}(\omega,l)\big)(\exists m \le \omega)(\forall k \le l)\varphi(n,m,k), \tag{2.103}$$

and the function $\overline{n}(k,l)$ is infinite for all l and infinite k. By corollary 2.57, there follows $(\forall l)(\forall^{st} n_1)(\exists^{st} n_2)(\overline{n}(n_2,l) > n_1)$, and, by the *uu*-principle, there holds

$$(\forall n_1)(\exists n_2)(\forall l)(\overline{n}(n_2,l) > n_1). \tag{2.104}$$

By definition (see (2.99)), $\overline{n}(n_2,l)$ is the largest $n' \le n_2$ such that

$$(\forall n \le n')(\exists m \le n_2)(\forall k \le l)\varphi(n,m,k).$$

This formula, together with (2.104), yields

$$(\forall n_1)(\exists n_2)(\forall l)\big(\forall n \le n_1\big)(\exists m \le n_2)(\forall k \le l)\varphi(n,m,k),$$

and in particular $(\forall n_1)(\exists n_2)(\forall l)(\exists m \le n_2)(\forall k \le l)\varphi(n_1,m,k)$. If we can prove that $(\forall l)(\exists m \le n_2)(\forall k \le l)\varphi(n,m,k)$ implies $(\exists m \le n_2)(\forall k)\varphi(n,m,k)$, then the right-hand side of (2.101) follows. The missing implication is the contraposition of an axiom of BERNA's replacement schema.

Second, we prove the reverse implication of (2.101). Assume the right-hand side of this formula holds, i.e. we have $(\forall n)(\exists m)(\forall k)\varphi(n,m,k)$. Fix $n_0 \in \mathbb{N}$ and Σ_2-transfer applied to $(\exists m)(\forall k)\varphi(n_0,m,k)$ yields $(\exists^{st} m)(\forall^{st} k)\varphi(n,m,k)$. This implies $(\forall^{st} n)(\exists^{st} m)(\forall^{st} k)\varphi(n,m,k)$ and we are done. □

2.63. Theorem. *In* BERNA*,* UUP *is equivalent to* Π_3-TRANS.

PROOF. By the previous theorem, UUP implies Π_3-TRANS.

Now assume Π_3-TRANS and let f be as in the *uu*-principle. Assume there holds $(\forall\vec{x})(\forall^{st}n)(\exists^{st}m)(f(m,\vec{x}) > n)$ and let ω be infinite. By corollary 1.55, there is a function $g(n,\vec{x})$ which calculates the least m such that $f(m,\vec{x}) > n$, for any $\vec{x}$ and $n \in \mathbb{N}$. Fix an infinite hypernatural ω_1 and define $h(n)$ as $\max_{\|\vec{x}\| \le \omega_1} g(n,\vec{x})$. This implies

$$(\forall^{st}n)(\exists m \le h(n))(\forall\vec{x})(\|\vec{x}\| \le \omega_1 \to f(m,\vec{x}) > n).$$

By noting that $h(n)$ is finite for $n \in \mathbb{N}$, we obtain

$$(\forall^{st}n)(\exists^{st}m)(\forall\vec{x})(\|\vec{x}\| \le \omega_1 \to f(m,\vec{x}) > n)$$

and also

$$(\forall^{st}n)(\exists^{st}m)(\forall^{st}\vec{x})(f(m,\vec{x}) > n).$$

By Π_3-transfer, this implies $(\forall n)(\exists m)(\forall\vec{x})(f(m,\vec{x}) > n)$ and we are done. □

Note that the proof fails if only $(\forall^{st}\vec{x})(\forall^{st}n)(\exists^{st}m)(f(m,\vec{x}) > n)$, i.e. it is not possible to pull a *standard* quantifier $(\forall^{st}\vec{x})$ through the quantifier $(\exists^{st}m)$.

Previously, we claimed that ideas, techniques and even proofs carry over between the stratified and classical framework. To see this, compare the proofs of theorems 2.62 and 2.63 with the proof of theorems 2.21 and 2.24.

The following generalizes UUP, without changing its essential character.

2.64. PRINCIPLE (General UUP). *For every quantifier-free $\varphi \in L^{st}$, we have*

$$(\forall\vec{x})(\forall^{st}n)(\exists^{st}m)\varphi(n,m,\vec{x}) \to (\forall n)(\exists m)(\forall\vec{x})\varphi(n,m,\vec{x}). \tag{2.105}$$

2.65. THEOREM. *In* BERNA*,* UUP *is equivalent to* General UUP*.*

PROOF. Analogous to the proofs of theorem 2.62 and 2.63. □

Heine's theorem states that every continuous function on a compact interval is uniformly continuous. Thus, it has the same syntactical form as the *uu*-principle: a universal quantifier is 'pulled through' an existential quantifier. The following theorem suggests a deeper connection.

2.66. THEOREM. *Let f be a standard function. In* BERNA + UUP

$$(\forall x \in [0,1])(\forall^{st}k)(\exists^{st}m)(\forall y \in [0,1])(|x-y| < 1/m \to |f(x)-f(y)| < \tfrac{1}{k}) \tag{2.106}$$

implies

$$(\forall^{st}k)(\exists^{st}m)(\forall x,y \in [0,1])(|x-y| < 1/m \to |f(x)-f(y)| < \tfrac{1}{k}). \tag{2.107}$$

PROOF. Let f be a standard function. Formula (2.106) implies

$$(\forall x \in [0,1])(\forall^{st}k)(\exists^{st}m)(\forall y \in [0,1])(\|y\| \le \|x\| \wedge |x-y| < \tfrac{1}{m} \to |f(x)-f(y)| < \tfrac{1}{k}),$$

and, by corollary 1.46, the subformula starting with '$(\forall y \in [0,1])$' may be treated as quantifier-free. Applying UUP to the previous formula yields

$$(\forall k)(\exists m)(\forall x,y \in [0,1])(\|y\| \le \|x\| \;\wedge\; |x-y| < \tfrac{1}{m} \to |f(x)-f(y)| < \tfrac{1}{k}),$$

and also

$$(\forall k)(\exists m)(\forall x,y \in [0,1])(|x-y| < \tfrac{1}{m} \to |f(x)-f(y)| < \tfrac{1}{k}).$$

Now fix $k_0 \in \mathbb{N}$ and let M be the least m such that

$$(\forall x,y \in [0,1])(|x-y| < \tfrac{1}{m} \to |f(x)-f(y)| < \tfrac{1}{k_0}).$$

By theorems 2.63 and 1.64, the number M is definable in BERNA + UUP. If M is infinite, then there are $x_0, y_0 \in [0,1]$ for which there holds $|x_0 - y_0| < \frac{1}{M-1}$ and $|f(x_0) - f(y_0)| \geq \frac{1}{k_0}$. But then $x_0 \approx y_0$ and applying (2.106) for $k = k_0$, $x = x_0$ and $y = y_0$ yields a contradiction. Thus, M is finite and we are done. □

Thus, UUP not only has the same 'syntactical' form as Heine's theorem, it also proves a nonstandard version of it. Moreover, Heine's theorem is equivalent to Weak König's Lemma over RCA_0, and WKL_0 has first-order strength of $I\Sigma_1$ (see [**46**]). Incidentally, BERNA + UUP is at least as strong as $I\Sigma_1$, by theorems 2.63 and 1.64. Thus, Π_3-TRANS is too strong for finitism, as $I\Sigma_1$ is stronger than PRA, although the former is Π_2-conservative over the latter. Also, theorem 2.63 is in agreement with Harvey Friedman's Grand Conjecture (see [**3**] and [**19**]):

> *Every theorem published in the Annals of Mathematics whose statement involves only finitary mathematical objects* (*i.e. what logicians call an arithmetical statement*) *can be proved in* EFA.

Indeed, although we used BERNA's replacement axioms, which go beyond $I\Delta_0 + \exp$, to prove theorem 2.63, the contents of this theorem, namely Π_3-TRANS, was not finitistic in nature to begin with.

To conclude this paragraph, we show another context in which BERNA comes up. Consider the following generalization of ATOM (see schema 1.156).

2.67. Axiom schema (MOL). *For every arithmetical f, if there is an $n \in \mathbb{N}$ such that $f(n, m)$ is infinite for all m, then there is a least number with this property.*

2.68. Theorem. *In* BERNA + Π_2-TRANS, Π_3^{st}-MIN *is equivalent to* MOL.

Proof. Analogous to the proof of theorem 1.159. □

Thus, BERNA repeatedly appears in the context of Reverse Mathematics.

5.2. Conservation and expansion for ERNA. In this section, we obtain a conservation result for ERNA. This, together with the results from the previous sections, yields a consistency proof for ERNA + Π_2-TRANS$^-$.

5.2.1. *Conservation for* ERNA+Π_1-TRANS$^-$. Here, we prove a Π_2-conservation result for ERNA. There are many such results in logic and equally so for nonstandard mathematics. Examples include [**2**], where Avigad and Helzner prove that a nonstandard version of PRA, extended with a transfer principle for Π_2-formulas, is conservative over PRA. There is [**32**], where Keisler proves that a nonstandard version of ACA_0 plus transfer for all arithmetical formulas is conservative over Peano arithmetic. Nonstandard set theories such as **IST**, **BST** and **GRIST** are -in a technical sense- conservative over **ZFC** (see [**23, 36, 39**]). However, in each case powerful techniques, like forcing, the Löwenheim-Skolem theorem or ultrapowers, are used to yield the conservation results. We wish to obtain a conservation result for ERNA that is both elementary in its techniques and provable in PRA or related systems of finitistic reductionism. In particular, we shall on occasion use Gödel's completeness theorem, available in WKL_0, and the latter is Π_2-conservative over PRA (see [**46**]).

As it turns out, our conservation result gives rise to a new consistency proof of ERNA + Π_2-TRANS$^-$. However, this consistency proof does not work without the above reformulation of Π_2-TRANS as the unboundedness principle.

Let ERNA^{st} be ERNA limited to axioms not involving the nonstandard objects ω, ε and $\approx$. Note that ERNA^{st} is essentially $I\Delta_0 + \exp$. We have the following conservation result.

2.69. Theorem (Π_2-conservation). *Let $\varphi \in L^{st}$ be quantifier-free. If* ERNA *proves* $(\forall^{st} n)(\exists^{st} m)\varphi(n,m)$*, then* ERNA^{st} *proves* $(\forall n)(\exists m)\varphi(n,m)$.

Proof. Before we start with the actual proof, recall that any natural number can be used for a_0 in (1.4), without affecting the correctness of the proof. In particular, we can replace (1.4) with

$$a_0 := n_0,\ b_0 := f_{D+1}(n_0),\ c_0 := b_0 \text{ and } d_0 := f_{D+1}(c_0), \tag{2.108}$$

for any nonzero $n_0 \in \mathbb{N}$, and we still obtain a valid consistency proof of ERNA, be it with larger numbers a_D, b_D, c_D and d_D.

Let φ be as in the theorem and assume that ERNA proves $(\forall^{st} n)(\exists^{st} m)\varphi(n,m)$. If ERNA^{st} does not prove $(\forall n)(\exists m)\varphi(n,m)$, then, by completeness, there is a model $\mathcal{M}$ of ERNA^{st} such that $\mathcal{M} \models (\exists n)(\forall m)\neg\varphi(n,m)$. Let c be a new constant. The following theory is consistent, for all $k \in \mathbb{N}$,

$$\mathrm{ERNA}^{st} + (\forall m)\neg\varphi(c,m) + (2^c_k > 0)\,. \tag{2.109}$$

By Herbrand's theorem, every finite set of instantiated axioms of this theory is consistent.

Now consider the theory 'ERNA $+ (\forall m)\neg\varphi(c,m) \wedge c$ is finite'. By Herbrand's theorem, it is consistent if every finite set T of instantiated axioms of this theory is consistent. Consider such a set T and let the number B be as in (1.2), i.e. we have $\|f(\vec{x})\| \leq 2^{\|\vec{x}\|}_B$ for all the functions (except min and c) appearing in T. Also, let D be the maximum depth of all terms in T. Then define $f_0(x)$ and $f_n(x)$ as is in (1.3) and assume k_1 is such that $f_{D+1}(f_{D+1}(x)) < 2^x_{k_1}$. Let T' be a finite set of instantiated axioms of (2.109), with $k = k_1$, which contains all ERNA^{st}-axioms of T, the axioms $(\forall m)\neg\varphi(c,m)$ of T and the axiom $2^c_{k_1} > 0$. By the above, we know that T' is consistent and hence has a model. Let val_1 be the interpretation, D_1 the domain and assume $\mathrm{val}_1(c) = n_1$.

Finally, we show how to adapt the interpretation val_1 in order to validly interpret T. The axioms of ERNA^{st} in T already have a valid interpretation. To interpret the other axioms, we perform the same D-step construction as in the consistency proof of ERNA with a_0 as in (2.108) and $n_0 = n_1$. By our choice of $k = k_1$, this construction takes place in the interval $[0, d_0]$. Indeed, by the previous, we have $d_0 = f_{D+1}(f_{D+1}(n_1)) < 2^{n_1}_{k_1}$ and hence the D-step construction certainly stays within the domain D_1, by the inclusion of '$2^c_{k_1} > 0$' in T'. Let val be the mapping obtained by this construction. Then all axioms of T, except for '$(\forall m)\neg\varphi(c,m) \wedge c$ is finite', have received a valid interpretation. Since we have that $a_0 \leq a_D$ and that 'val(τ is finite)' is equivalent to $|\tau| \leq a_D$, our choice of $a_0 := n_1$ guarantees that c is interpreted as a finite number. Hence, the axiom '$(\forall m)\neg\varphi(c,m) \wedge c$ is finite' is interpreted as true and we have obtained a valid interpretation of T. Hence, the theory 'ERNA $+ (\forall m)\neg\varphi(c,m) \wedge c$ is finite' is consistent, but his contradicts the assumption that ERNA $\models (\forall^{st} n)(\exists^{st} m)\varphi(n,m)$. Hence, ERNA^{st} does prove $(\forall n)(\exists m)\varphi(n,m)$ and we are done. □

2.70. Corollary (Π_2-conservation). *Let φ be as in the theorem. If* ERNA + Π_1-TRANS$^-$ *proves* $(\forall^{st} n)(\exists^{st} m)\varphi(n,m)$*, then* ERNAst *proves* $(\forall n)(\exists m)\varphi(n,m)$.

Proof. In the proof of the previous theorem, the model for the subset T of the theory 'ERNA + $(\forall m)\neg\varphi(c,m) \wedge c$ is finite' can be expanded to a model of 'ERNA + Π_1-TRANS + $(\forall m)\neg\varphi(c,m) \wedge c$ is finite ', using the same techniques as used in theorem 1.58, by choosing k_1 large enough to accommodate the various iterations of ERNA's consistency proof required for the consistency proof of ERNA + Π_1-TRANS$^-$. □

5.2.2. *The consistency of* ERNA+Π_2-TRANS$^-$. Here, we combine results from the previous sections to obtain a new consistency proof of the theory ERNA + Π_2-TRANS$^-$, introduced in [**29**]. Both the unboundedness principle and the above conservation result are crucial for this proof. For brevity, we write Φ_f for '$(\forall^{st} n)(\exists^{st} m)(f(m) > n) \rightarrow (\forall n)(\exists m)(f(m) > n)$'. We have the following theorem.

2.71. Theorem. *For $1 \leq i \leq N$, let $f_i \in L^{st}$ be weakly increasing non-negative functions. Then* ERNA + $\bigcup_{i=1}^{N} \Phi_{f_i}$ *is consistent.*

Proof. First, we treat the case $N = 1$. Let $f = f_1$ be as in the theorem. Assume to the contrary that ERNA + Φ_f is inconsistent. Then, in every model of ERNA, there holds $\neg\Phi_f$, i.e.

$$(\forall^{st} n)(\exists^{st} m)(f(m) > n) \wedge (\exists n)(\forall m)(f(m) \leq n). \tag{2.110}$$

By completeness, ERNA proves $(\forall^{st} n)(\exists^{st} m)(f(m) > n)$. By theorem 2.69, ERNAst proves $(\forall n)(\exists m)(f(m) > n)$. Thus, the latter holds in every model of ERNAst. As a model for ERNA is also a model of ERNAst, the previous contradicts the second part of the conjunction (2.110). Thus, the theory ERNA + Φ_f is consistent.

Assume that the theorem holds for $N \geq 1$. We show that it holds for $N+1$. Let $f_i \in L^{st}$, for $1 \leq i \leq N+1$, be as in the theorem. Assume to the contrary that ERNA + $\cup_{i=1}^{N+1}\Phi_{f_i}$ is inconsistent. Then, in every model of ERNA, there holds $\bigvee_{i=1}^{N+1} \neg\Phi_{f_i}$, i.e.

$$\bigvee_{i=1}^{N+1} \big[(\forall^{st} n)(\exists^{st} m)(f_i(m) > n) \wedge (\exists n)(\forall m)(f_i(m) \leq n)\big] \tag{2.111}$$

Define $g(m) = \sum_{i=1}^{N+1} f_i(m)$. In every model $\mathcal{M}$ of ERNA, we have the formula $(\forall^{st} n)(\exists^{st} m)(g(m) > n)$, regardless of which part of the disjunction (2.111) is true in $\mathcal{M}$. By completeness, ERNA proves $(\forall^{st} n)(\exists^{st} m)(g(m) > n)$. By theorem 2.69, ERNAst proves $(\forall n)(\exists m)(g(m) > n)$. Moreover, ERNAst proves $(\forall n)(\exists m)(g(m) > (N{+}1)n)$ and by Parikh's theorem (see [**8**]), also $(\forall n)(\exists m \leq 2^n_{k_0})\big(g(m) > (N{+}1)n\big)$, for some $k_0 \in \mathbb{N}$. As all f_i are weakly increasing, this implies $(\forall n)(g(2^n_{k_0}) > (N+1)n)$. Finally, this yields $(\forall m)(g(m) > (N+1)\log^{k_1} m)$, for some $k_1 \in \mathbb{N}$, in ERNAst. As ERNAst is a subset of ERNA, the latter also proves $(\forall m)(g(m) > (N+1)\log^{k_1} m)$.

Suppose that for $1 \leq i \leq N+1$, the functions f_i are eventually dominated by all functions $\log^k x$ for $k \in \mathbb{N}$. This means that, for all $k \in \mathbb{N}$,

$$(\forall i \in \{1,\dots,N+1\})(\exists K)(\forall m \geq K)(f_i(m) \leq \log^k(m)). \tag{2.112}$$

Using Π_1-REPL, this implies, for all $k \in \mathbb{N}$, that

$$(\exists L)(\forall i \in \{1, \ldots, N+1\})(\exists K \leq L)(\forall m \geq K)(f_i(m) \leq \log^k(m)). \tag{2.113}$$

This immediately yields, for all $k \in \mathbb{N}$, that

$$(\exists K)(\forall i \in \{1, \ldots, N+1\})(\forall m \geq K)(f_i(m) \leq \log^k(m)).$$

Summing all f_i in the previous formula, we obtain, for all $k \in \mathbb{N}$, that

$$(\exists K)(\forall m \geq K)(f_1(m) + \cdots + f_{N+1}(m) \leq (N+1)\log^k(m)).$$

But we previously proved that $(\forall m)(g(m) > (N+1)\log^{k_1} m)$. Thus, we have a contradiction and one of the functions f_i is not eventually dominated by all functions $\log^k x$. Suppose it is f_{i_0}. Hence, there is a $k_2 \in \mathbb{N}$ such that

$$(\forall K)(\exists m \geq K)(f_{i_0}(m) > \log^{k_2}(m)).$$

Thus, $f_{i_0}(n)$ has growth rate similar to $\log^{k_2}(n)$ and hence the following formula is false in any model of ERNA

$$(\forall^{st} n)(\exists^{st} m)(f_{i_0}(m) > n) \wedge (\exists n)(\forall m)(f_{i_0}(m) \leq n).$$

This implies that (2.111) is equivalent to

$$\bigvee_{i=1, i\neq i_0}^{N+1} \big((\forall^{st} n)(\exists^{st} m)(f_i(m) > n) \wedge (\exists n)(\forall m)(f_i(m) \leq n)\big).$$

But by assumption, $\bigcup_{i=1,i\neq i_0}^{N+1} \Phi_{f_i}$ is consistent with ERNA, which yields a contradiction. Thus, ERNA $+ \bigcup_{i=1}^{N+1} \Phi_{f_i}$ is consistent. $\square$

2.72. Corollary. *The theory* ERNA $+ \Pi_2$-TRANS$^-$ *is consistent.*

Proof. Immediate from the compactness theorem and corollary 2.59. $\square$

The proof of the previous theorem hinges on the fact that the functions f_i eventually dominate $\log^k x$ for some fixed $k \in \mathbb{N}$. This seems arbitrary as ERNA can define even slower growing functions, like e.g. $\log^* x$ (see paragraph 3.2.4). In section 5.3, we show why the iterations of the log-function are essential to the proof of theorem 2.71. In fact, we show that these functions are fundamental to the very concept of provability in ERNA.

In the following final paragraphs of this section, we discuss the connection between mathematical practice and Π_2-transfer.

First of all, it is easy to show that the inverse implication in (2.98) in Π_2-TRANS may be omitted. Thus, every axiom in the latter schema may be assumed to be of the form

$$(\forall^{st} n)(\exists^{st} m)\varphi(n,m) \to (\forall n)(\exists m)\varphi(n,m). \tag{2.114}$$

In practice, we would only use (2.114) together with $(\forall^{st} n)(\exists^{st} m)\varphi(n,m)$ to conclude $(\forall n)(\exists m)\varphi(n,m)$, by modus ponens.

Now, let $\psi \in L^{st}$ be quantifier-free and assume ERNA $+ \Pi_2$-TRANS proves the formula $(\forall^{st} n)(\exists^{st} m)\psi(n,m)$ with proof P. Additionally assume that Π_2-transfer is only used in the way described in the previous paragraph, i.e. if (2.114) occurs in P, the latter must include a proof P' of $(\forall^{st} n)(\exists^{st} m)\varphi(n,m)$ and together with (2.114), it is concluded in P that $(\forall n)(\exists m)\varphi(n,m)$, by modus ponens. If P' is an ERNA-proof, then ERNAst proves $(\forall n)(\exists m)\varphi(n,m)$, by theorem 2.69. Hence, the modus ponens inference that yields $(\forall n)(\exists m)\varphi(n,m)$ may be replaced by an ERNAst-proof. In this way, we can remove all occurrences of (2.114) from P, starting at the top of the proof. Thus, we obtain an ERNA-proof of $(\forall^{st} n)(\exists^{st} m)\psi(n,m)$ and by theorem 2.69, ERNAst proves $(\forall n)(\exists m)\psi(n,m)$.

Thus, there seems to be a significant difference between mathematical practice and logical strength. Indeed, by theorem 1.75, ERNA + Π_2-TRANS is much stronger than ERNA, although they prove the same Π_2-statements if we limit the use of Π_2-transfer to that of 'ordinary' mathematics. This can be considered as evidence for Friedman's Grand Conjecture (see the quote before schema 2.67 and [**3**, **19**]).

Finally, we suspect that the argument of the previous paragraphs can be used to prove that a nonstandard extension of PRA plus Π_2-transfer is Π_2-conservative over PRA ([**1**]).

5.3. Intensionality. In this section, we explore the connection between provability in ERNA and properties of the iterations of the log-function.

5.3.1. *Intensional objects.* The following definitions are crucial. They supersede any previous definitions.

2.73. Definition. A term $\tau(n)$ is called 'arithmetical' if it is in L^{st}, non-negative, weakly increasing in n and does not involve min.

A k-ary term is arithmetical if it is arithmetical in every variable.

2.74. Definition. An arithmetical term $\tau(n)$ is called 'intensional' if there is a $k \in \mathbb{N}$ such that $\tau(n)$ eventually dominates $\log^k n$ for $n \in \mathbb{N}$.

The best-known example of a 'non-intensional' function is $\log^* x$ (see paragraph 3.2.4). Indeed, it grows slower than $\log^k x$ for all $k \in \mathbb{N}$ and for $n_0 = 2^{65536}$, which is larger than the number of particles in the universe, $\log^* n_0$ is at most five. Thus, for practical purposes, $\log^* x$ may be regarded as a constant function, although PRA (or $I\Delta_0$ + superexp) proves that it is unbounded. The following theorem makes this qualitative statement precise and more convincing.

2.75. Theorem. *The theory $I\Delta_0 + \exp$ cannot prove that $\log^* x$ is unbounded, i.e.* $I\Delta_0 + \exp \nvdash (\forall x)(\exists y)(\log^* y > x)$.

Proof. Assume to the contrary that $I\Delta_0 + \exp$ proves $(\forall x)(\exists y)(\log^* y > x)$. By Parikh's theorem (see [**8**]), there is a term t such that $I\Delta_0$+exp proves $(\forall x)(\exists y \leq t(x))(\log^* y > x)$. As $\log^* x$ is weakly increasing, there follows $(\forall x)(\log^*(t(x)) > x)$. However, this implies that $t(x)$ grows faster than all 2^x_k, which is impossible. $\square$

By completeness, there is a model of $I\Delta_0 + \exp$ in which $\log^* x$ is bounded. From the point of view of logic, this model is 'nonstandard' and 'exotic'. However, given the slow-growing nature of $\log^* x$ discussed above, we perceive this function as eventually constant in the 'real world'. Thus, this 'exceptional' model is natural from the anthropocentric point of view. This is the idea behind the proof of the Isomorphism Theorem (see paragraph 3.2.4).

In the same way as in the theorem, one can show that PRA does not prove the unboundedness of $A^{-1}(x)$, the inverse of the well-known Ackermann function. Again, by completeness, there is a model of PRA in which $A^{-1}(x)$ is bounded. Thus, there is also a model of $I\Delta_0$+exp in which $A^{-1}(x)$ is bounded and $\log^*(x)$ is unbounded.

Furthermore, there exist models of $I\Delta_0 + \exp$ in which an arbitrary non-intensional function is unbounded, but slower growing function are bounded. In this way, the iterations of the log-function are the 'resolution' of $I\Delta_0 + \exp$: this theory cannot 'detect finer objects', i.e. slower growing functions cannot be distinguished from the constant functions.

2.76. Corollary. *The theory* ERNA *cannot prove that the function* $\log^* x$ *is unbounded on* $\mathbb{N}$, *i.e.* ERNA $\not\vdash (\forall^{st} x)(\exists^{st} y)(\log^* y > x)$.

Proof. Immediate from theorem 2.69 and the fact that ERNA^{st} is essentially $I\Delta_0 + \exp$. $\square$

The following corollary shows that the nonstandard framework yields very elegant quantifier-free unprovable statements.

2.77. Corollary. *The theory* ERNA *cannot prove that* $\log^* \omega$ *is infinite.*

Proof. We show that even ERNA + Π_1-TRANS cannot prove that $\log^* \omega$ is infinite. Assume to the contrary that ERNA + Π_1-TRANS does prove that $\log^* \omega$ is infinite. This implies $(\forall^{st} n)(\log^* \omega > n)$ and also $(\forall^{st} n)(\exists m)(\log^* m > n)$. Applying Σ_1-transfer to the latter formula yields that ERNA + Π_1-TRANS proves $(\forall^{st} n)(\exists^{st} m)(\log^* m > n)$. By theorem 2.69, ERNA^{st} proves $(\forall x)(\exists y)(\log^* y > x)$, which contradicts theorem 2.75 and we are done. $\square$

2.78. Corollary. *Let* f *be non-intensional. The theory* $I\Delta_0 + \exp$ *cannot prove that* $f(x)$ *is unbounded, i.e.* $I\Delta_0 + \exp \not\vdash (\forall x)(\exists y)(f(y) > x)$.

Proof. Immediate from the proof of the theorem and definition 2.74. $\square$

Besides $\log^* x$, there are quite a number of non-intensional objects in $I\Delta_0 + \exp$. Indeed, if $f(x)$ is non-intensional, then $g(x) := (\mu m \leq x)(f^m(x) \leq 1)$ grows slower than all iterations of $f(x)$. If we iterate this minimization procedure enough, we obtain a function which grows slower than the inverse of any fixed primitive recursive function. Thus, $I\Delta_0 + \exp$ contains a 'mirrored copy' of the Grzegorczyk hierarchy of PRA. However, given the above, $I\Delta_0 + \exp$ cannot prove anything about this copy. Using diagonalization, we can obtain even slower growing functions in $I\Delta_0 + \exp$, like e.g. the inverse Ackermann function.

Theorem 2.75 shows that ERNA cannot prove elementary properties concerning unboundedness and infinitude of non-intensional functions. This is a good point to explain our use of the word 'intensional'. In order to prove Gödel's famous incompleteness theorems, the syntax of first-order logic is coded with integers. This process is called the 'arithmetization' of metamathematics (see [**8**, Chapter II] for details) and it can be done essentially in two ways: via numeralwise representation or via the intensional approach. In the latter, the syntactical concepts of a theory T such as 'theorem' or 'formula' are not merely coded into terms, but the theory T can prove simple properties of these terms. For instance, in the intensional approach, T proves that the set of T-theorems is closed under modus ponens. Since $I\Delta_0 + \exp$ cannot even prove that $\log^* x$ is unbounded, it seems appropriate to call this function 'non-intensional'. For more details regarding intensionality, see [**8**, p. 113] and [**15**].

We introduced overflow in theorem 2.55. The following definition makes the dependence of $\overline{n}$ on φ in corollary 2.56 more apparent.

2.79. Definition. For a quantifier-free formula $\varphi \in L^{st}$, we define

$$\overline{n}_\varphi(k) := (\mu n \leq k)(\forall m \leq k)\neg\varphi(n+1, m). \tag{2.115}$$

The notion of intensional object gives rise to the following theorem. It is analogous to Parikh's theorem (see [**8**]).

2.80. Theorem. *Assume $\varphi \in L^{st}$ is quantifier-free. Then $(\forall^{st} n)(\exists^{st} m)\varphi(n, m)$ is provable in* ERNA *if and only if $\overline{n}_\varphi$ is intensional.*

Proof. Assume ERNA proves $(\forall^{st} n)(\exists^{st} m)\varphi(n, m)$. For fixed infinite k_0, this implies $(\forall^{st} n)(\exists m \leq k_0)\varphi(n, m)$ and overflow yields

$$(\forall n \leq \overline{n}_\varphi(k_0))(\exists m \leq k_0)\varphi(n, m).$$

This proves that $\overline{n}_\varphi(k)$ is infinite for infinite k. By corollary 2.57, ERNA also proves $(\forall^{st} m)(\exists^{st} k)(\overline{n}_\varphi(k) > m)$. By theorem 2.69, $I\Delta_0 + \exp$ proves $(\forall m)(\exists k)(\overline{n}_\varphi(k) > m)$. By corollary 2.78, the function $\overline{n}_\varphi(k)$ must be intensional.

Assume that $\overline{n}_\varphi(k)$ is intensional. Thus, there is a $k_0 \in \mathbb{N}$ such that $(\forall^{st} m \geq k_0)(\overline{n}_\varphi(m) > \log^{k_0} m)$. Since ERNA proves the unboundedness of $\log^{k_0}(m)$ on $\mathbb{N}$, the previous implies the unboundedness of $\overline{n}_\varphi(m)$ on $\mathbb{N}$. Thus, there holds $(\forall^{st} k')(\exists^{st} k)\big[\overline{n}_\varphi(k) > k'\big]$ and as $\overline{n}_\varphi(k)$ is the largest $n' \leq k$ such that $(\forall n \leq n')(\exists m \leq k)\varphi(n, m)$, this yields $(\forall^{st} n)(\exists^{st} m)\varphi(n, m)$ and we are done. □

Thus, we showed that provability in ERNA is intimately connected with the iterations of the log-function. Also, the previous implies that it is undecidable whether a function is intensional or not.

2.81. Corollary. *Assume $\varphi \in L^{st}$ is quantifier-free. The formula $(\forall n)(\exists m)\varphi(n, m)$ is provable in $I\Delta_0 + \exp$ if and only if $\overline{n}_\varphi$ is intensional if and only if $\overline{n}_\varphi(\omega)$ is provably infinite in* ERNA*.*

Proof. Immediate from theorem 2.69 and the fact that ERNAst is essentially $I\Delta_0 + \exp$. □

5.3.2. *Phase transitions.* We now consider an interesting phenomenon which can occur when the formula φ in the previous theorem depends on a functional parameter f.

Assume that $I\Delta_0 + \exp$ proves $(\forall n)(\exists m)\varphi(n, m, f)$ for f growing faster than a certain function f_0 and does not prove $(\forall n)(\exists m)\varphi(n, m, f)$ for f growing slower than f_0. We say that $(\forall n)(\exists m)\varphi(n, m, f)$ experiences a 'phase transition' (from provability to unprovability) at f_0. By the previous corollary, the function $\overline{n}_\varphi(k)$ is intensional if and only if f grows faster than f_0. Thus, when f varies from faster growing than f_0 to slower growing than f_0, the formula $(\forall n)(\exists m)\varphi(n, m, f)$ goes from provable to unprovable (in $I\Delta_0 + \exp$) and the function $\overline{n}_\varphi(k)$ goes from a $\log^k x$ growth rate to a $\log^* x$ growth rate (or slower). Thus, every phase transition in $I\Delta_0 + \exp$ corresponds to a change in growth rate from $\log^k x$ to $\log^* x$ (or slower). By theorem 2.75, the latter growth rate change can be seen as 'the simplest phase transition' from provability to unprovability. Thus, we have showed that all phase transitions for Π_2-formulas are of this simple form.

Another interesting fact concerns the 'threshold' function f_0. By the above, we know that $I\Delta_0 + \exp$ cannot distinguish between functions which grow slower than all iterations of the log-function. Thus, the 'finest' variation of a function parameter that is available in $I\Delta_0 + \exp$ is from iterations of the log functions to $\log^* x$, as the latter is bounded above in some models, i.e. essentially a constant. Hence, the 'sharpest' phase transition we can expect to obtain in $I\Delta_0 + \exp$ will involve $\log^* x$ and $\log^k x$. We have the following conjecture.

2.82. CONJECTURE. *Let $f(n, d)$ be weakly increasing in d and arithmetical and non-intensional for fixed $d \in \mathbb{N}$. Then $f(n, \log^*(n))$ is non-intensional too.*

The condition that f be weakly increasing in d, stems from the fact that in practice, the inverse of f is used and this inverse is decreasing in d (see [**55**]).

We have the following partial sketch of a proof for the conjecture. By corollary 2.78, for each non-intensional function g, there is a model of $I\Delta_0 + \exp$ in which g is bounded above. Using properties of the Grzegorczyk hierarchy of PRA, it is not difficult to show that for a finite set of non-intensional functions, there is a model in which all these functions are bounded above. By compactness, there is such a model for a countable set of non-intensional functions. Let f be as in the conjecture. Then there is a model $\mathcal{M}$ such that $\log^*(n)$ and $f(n, k_0)$, for each $k_0 \in \mathbb{N}$, are bounded above. In this model, $f(n, \log^* n)$ is below $f(n, k_1)$ for some $k_1 \in \mathbb{N}$. Thus, $f(n, \log^* n)$ is bounded in $\mathcal{M}$ and hence non-intensional.

A nonstandard proof of the above conjecture could involve Herbrand's theorem instead of the compactness theorem, since the statement '$g(\omega)$ is infinite' is quantifier-free.

5.3.3. *Generalizations.* All of the above is easily generalized to stronger theories. Indeed, a theory of arithmetic T proves that $H_\alpha^{-1}(x)$ is unbounded, if and only if $\alpha < |T|$, where $|T|$ is T's proof-theoretical ordinal (see [**8**, Chapter III] for details). Hence, a function is called 'intensional' in T, if it eventually dominates $H_\alpha^{-1}(x)$ for some $\alpha < |T|$. In particular, the 'sharpest' phase transitions we can expect in T will always involve $H_{|T|}^{-1}(x)$. Also, let ${}^*T$ be a nonstandard conservative extension of T. Then ${}^*T$ proves that $H_\alpha^{-1}(\omega)$ is infinite if and only if $\alpha < |T|$. Thus, the stronger a theory T, the larger its proof-theoretic ordinal $|T|$ and the 'more' provably infinite numbers of the form $H_\alpha^{-1}(\omega)$ there are in ${}^*T$, i.e. the 'longer' the segment of provably infinite numbers below ω is. This corresponds to the intuition that stronger theories prove the well-ordering of 'longer' well-orderings, i.e. those with larger order type.

Similarly, we can consider weaker theories of bounded arithmetic like S_2 and its fragments (see [**8**] for details). In this setting, the log function is non-intensional, i.e. S_2 cannot prove that $|x| := \lceil \log_2(x+1) \rceil$ is unbounded. This fact can be used in the following indirect way. Let T be a theorem (of Theoretical Computer Science) which involves bounding terms like $(s + |e|)^{O(1)}$. If S_2 proves T, then let $\mathcal{M}$ be a model of S_2 in which $|x|$ is bounded. In $\mathcal{M}$, the term $(s + |e|)^{O(1)}$ reduces to $s^{O(1)}$ and hence T also holds (in $\mathcal{M}$) for the better bound $s^{O(1)}$. Thus, S_2 cannot disprove the stronger version of T.

6. Concluding remarks

In the previous section, we showed that techniques, proofs and methods, carry over from the classical to the stratified framework, and vice versa. To conclude, we formulate some philosophical considerations concerning ERNA$^{\mathbb{A}}$.

First, we observe that, from a finitistic point of view, Π_1^β, Π_2^β and Π_3^β-transfer are preferable to Π_1^α, Π_2^α and Π_3^α-transfer, as the former transfer principles only refer to certain levels of infinity, whereas the latter refer to the totality of numbers. This observation is especially true when dealing with analysis. Indeed, in this chapter we proved that basic analysis can be obtained in an elegant and quantifier-free way in ERNA$^{\mathbb{A}}$ extended with Stratified Transfer. In particular, for the purpose of analysis, five degrees of infinity seem to suffice and Stratified Transfer limited to a discrete number of levels seems finitistically acceptable.

Furthermore, we have obtained elegant equivalent versions of several transfer principles of both ERNA and ERNA$^{\mathbb{A}}$. However, comparing corollary 2.25 and theorem 2.63, we notice a discrepancy: Π_1^α-transfer appears in the former, but Π_1-transfer does not appear in the latter. This is because (Weak) Stratified Transfer does not imply Π_1^α-TRANS, in general. In order to remedy this, we can add to ERNA$^{\mathbb{A}}$ an axiom '$\mathbb{A}$LL' stating that every number is α-finite for some $\alpha \in \mathbb{A}$ and β-infinite for $\beta \prec \alpha$. In the terminology of [**24**], $\mathbb{A}$LL states that every number has a minimal context level. Also, $\mathbb{A}$LL implies that there are no numbers larger than all ω_α ($\alpha \in \mathbb{A}$), i.e. that the latter numbers are 'all there is'. In ERNA$^{\mathbb{A}}+\mathbb{A}$LL, the (very) Weak Stratified Transfer principle implies Π_2^α-transfer, which yields Π_1^α-TRANS, in the same way as in the proof of theorem 2.58. In this way, corollary 2.25 simplifies in the extended theory ERNA$^{\mathbb{A}}+\mathbb{A}$LL.

However, the axiom $\mathbb{A}$LL has its problems. First of all, it conflicts with ERNA$^{\mathbb{A}}$'s quantifier-free and finitistic nature. Indeed, first of all, $\mathbb{A}$LL cannot easily be written as a quantifier-free formula. Second, $\mathbb{A}$LL implicitly refers to the totality of all numbers, which is not compatible with finitism as understood by Tait ([**51**]). Third, in ERNA$+\mathbb{A}$LL, Π_3^β-transfer implies Π_3^α-transfer, which makes the resulting theory at least as strong as $I\Sigma_1$, by theorem 1.64 and theorem 3.10 suggests that this theory is as strong as $I\Sigma_3$. Fourth, the following (meta)theorem shows that $\mathbb{A}$LL has peculiar properties.

2.83. THEOREM. *Let $\mathbb{A}$ be dense and infinite. In* ERNA$^{\mathbb{A}}+\mathbb{A}$LL*, there is a sequence of levels which satisfies each instance* (2.83) *of Π_1^β-transfer.*

PROOF. Assume $\alpha \succeq \mathbf{0}$ and consider $(\forall^{\alpha\text{-}st} n)\varphi(n)$ as in (2.83). Apply overflow to obtain $(\forall n \leq \overline{n})\varphi(n)$. Then $\overline{n}$ is β-finite for some $\beta \succ \alpha$ and we have $(\forall^{\gamma\text{-}st} n)\varphi(n)$ for $\alpha \prec \gamma \prec \beta$. As $\mathbb{A}$ is dense, the theorem follows. □

From the proof, it is clear that the same holds for every finite set of instances.

2.84. COROLLARY. *Let $\mathbb{A}$ be dense and infinite. In* ERNA$^{\mathbb{A}}+\mathbb{A}$LL$+\Pi_n^\beta$-TRANS*, there is a sequence of levels which satisfies each instance of Π_{n+1}^β-transfer.*

PROOF. We prove the theorem for $n = 2$. The general case follows from the particular case by ERNA$^{\mathbb{A}}+\mathbb{A}$LL's version of theorem 3.8. Fix $\alpha \succ \mathbf{0}$ and consider $(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)\varphi(n,m)$ as in (2.92). For $\beta \succ \alpha$, we have $(\forall^{\alpha\text{-}st} n)(\exists m \leq \omega_\beta)\varphi(n,m)$. Applying overflow yields $(\forall n \leq \overline{n})(\exists m \leq \omega_\beta)\varphi(n,m)$ where $\overline{n}$ is δ-infinite for some $\delta \succ \alpha$ and we have $(\forall^{\gamma\text{-}st} n)(\exists^{\beta\text{-}st} m)\varphi(n,m)$ for $\alpha \prec \gamma \prec \delta$. Applying Σ_1^β-transfer finishes the proof. □

Thus, the axiom $\mathbb{A}$LL clearly has its problems. An alternative solution which avoids $\mathbb{A}$LL, is to bound each instance of the universal quantifier $(\forall \vec{x})$ in the Weak Stratified Transfer Principle to the δ-finite numbers for $\delta \succ \beta \succ \alpha$. Then, the universal quantifier $(\forall l)$ in (2.89) can be replaced with $(\forall^{\delta\text{-}st} l)$ and the rest of the proof goes through unchanged. Indeed, we can easily obtain $(\forall^{\delta\text{-}st} l)(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(\forall k \leq l)\varphi(n,m,k)$ from $(\forall^{\alpha\text{-}st} n)(\exists^{\alpha\text{-}st} m)(\forall^{\alpha\text{-}st} k)\varphi(n,m,k)$ using Π_1^β-transfer, which is implied by Π_2^β-transfer. However, things become more cluttered this way, and it seems we have to make a choice between philosophy and aesthetics.

2.85. Acknowledgement. I wish to thank Karel Hrbacek and Andreas Weiermann for their valuable advice with regard to this chapter. The results in the first four sections of this chapter are published in [**44**].

CHAPTER III

Relative arithmetic

The name that can be named, is
not the eternal Name.

Tao Te Ching
LAO TSE

Introduction: internal beauty

Every mathematician is well-acquainted with the sequence $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. As a student, one witnesses the successive extensions, starting with $\mathbb{N}$ and finally ending up in $\mathbb{C}$. In Nonstandard Analysis, the real numbers are enriched with infinitesimals and their inverses to form $^*\mathbb{R}$, the hyperreal numbers. In this way, the nonstandard framework is presented as a logical next step in the above expansion process. Indeed, one creates a 'larger' set with a richer structure which allows for a uniform and elegant treatment of the problems in the original set. This approach to Nonstandard Analysis is called the 'external' viewpoint and was pioneered by Abraham Robinson in his milestone book [**41**].

Edward Nelson later introduced an alternative theory of Nonstandard Analysis called **IST** in [**36**]. Instead of extending the universe of objects, he merely adds a predicate 'x is standard' and asserts that there are standard and nonstandard objects. The properties of the new predicate are described in the three axiom schemas Idealization, Standardization and Transfer. We stress that no new objects are introduced: the new predicate just gives more structure to the universe of objects. Thus, Nelson's approach to Nonstandard Analysis is called the 'internal' viewpoint. This approach is far more elegant than the external one and obviously more ontologically parsimonious. Nelson's 'virtue of simplicity' ([**38**]) is embodied in his internal viewpoint. However, there is an inevitable tradeoff for this elegance and simplicity. Indeed, the so-called 'illegal set formation rule' prohibits the existence of the set $\{x \in \mathbb{N} \mid x \text{ is standard}\}$ in **IST**. Thus, the 'set of all standard naturals' is not available in **IST** and this seems strange compared to the external viewpoint, where both $\mathbb{N}$ and $^*\mathbb{N}$ are available. This asymmetry is an essential ingredient of **IST** and, as we shall see later, also of Nelson's philosophical views.

In this chapter, we study stratified nonstandard arithmetic from the internal viewpoint. In this way, we do not need to specify a set of levels $\mathbb{A}$ up front and no new objects are introduced. We only define a new predicate $x \sqsubseteq y$ with properties described in axiom schema 3.1. We sometimes write 'x is y-finite' instead of $x \sqsubseteq y$. This notation is purely symbolic and we may also read $x \sqsubseteq y$ as e.g. 'x is not very large compared to y'. The reader should verify that the axiom schema NS satisfies the intuitive laws that govern the notion of largeness in the 'real world' (or any similarly vague concept).

After fixing the basic axioms in NS, we introduce the classical transfer principle in our system. Using the latter, we can prove the 'reduction theorem' (see theorem 3.8) which reduces any arithmetical sentence to an equivalent Δ_0-sentence. Thus, it is possible to collapse the arithmetical hierarchy onto Δ_0, yielding a new link between Peano arithmetic and bounded arithmetic. Surprisingly, the reduction theorem is also equivalent to the aforementioned transfer principle (see theorem 3.15). As applications, we define a truth definition for arithmetical sentences and we formalize Nelson's notion of impredicativity (see [**37**]).

1. Internal relativity

In this section, we describe stratified nonstandard arithmetic and its fundamental features. Let L be the language of arithmetic. We introduce a new binary predicate '$x \sqsubseteq y$' which applies to all natural numbers. For better readability we write 'x is y-finite' instead of $x \sqsubseteq y$. The following axiom set describes the properties of $x \sqsubseteq y$. These axioms are not intended to be minimal.

3.1. Axiom schema (NS).

(1) *The numbers* 0, 1 *and* x *are* x*-finite.*
(2) *If* x *and* y *are* z*-finite, so are* $x + y$ *and* $x \times y$*.*
(3) *If* x *is* y*-finite and* $z \leq x$*, then* z *is* y*-finite.*
(4) *If* x *if* y*-finite and* y *is* z*-finite, then* x *is* z*-finite.*
(5) *Either* x *is* y*-finite or* y *is* x*-finite.*
(6) *There is a number* y *that is not* x*-finite.*

3.2. Definition. A number y is called 'x-infinite' if it is not x-finite. We denote this by '$x \ll y$'. A number is also called 'x-standard' if it is x-finite.

By item (6) of the previous schema, the set of natural numbers is 'stratified' in different 'levels' or 'degrees' of magnitude. Intuitively, numbers of the same level are 'finite' (or 'not very large') relative to each other and 'infinite' (or 'very large') compared to numbers of lower levels. The numbers 0 and 1 are at the lowest level.

It should be stressed that we do *not* expand the set of natural numbers; we only define a new predicate $x \sqsubseteq y$ which can be interpreted in several ways (see also section 6).

3.3. Definition. A formula is called 'internal' if it does not involve the predicate 'x is y-finite' for any x and y. Non-internal formulas are called 'external'.

In the following, we assume that the classes Δ_0, Σ_n and Π_n of the arithmetical hierarchy are limited to internal formulas, i.e. they carry their usual meaning. We also assume that all parameters are shown, unless explicitly stated otherwise.

3.4. Notation. We write '$(\exists^{x\text{-}st} y)\varphi(y)$' instead of $(\exists y)(y \text{ is } x\text{-finite} \wedge \varphi(y))$ and we write '$(\forall^{x\text{-}st} y)\varphi(y)$' instead of $(\forall y)(y \text{ is } x\text{-finite} \rightarrow \varphi(y))$.

Now consider the following transfer principle.

3.5. Axiom schema (Σ_n-TRANS). *For every formula* $\varphi \in \Delta_0$ *and* x*-finite* $\vec{y}$*,*

$$(\exists x_1)(\forall x_2)\ldots(Qx_n)\varphi(x_1, \ldots, x_n, \vec{y}) \tag{3.116}$$

is equivalent to

$$(\exists^{x\text{-}st}x_1)(\forall^{x\text{-}st}x_2)\dots(Q^{x\text{-}st}x_n)\varphi(x_1,\dots,x_n,\vec{y}). \tag{3.117}$$

Depending on whether n is odd or even, '(Qx_n)' is '$(\exists x_n)$' or '$(\forall x_n)$'.

For fixed x and $\varphi \in \Delta_0$, the previous schema is just the usual transfer principle for Σ_n-formulas, relative to the level of magnitude of x. Thus, Σ_n-TRANS expresses Leibniz's principle that the same laws should hold for all numbers, standard or nonstandard alike, relative to the level at which the numbers occur. For brevity, we write 'TRANS' for '$\cup_{n\in\mathbb{N}}\Sigma_n$-TRANS'.

By contraposition, the schema Σ_n-TRANS immediately yields the following equivalent transfer principle.

3.6. Axiom schema (Π_n-TRANS). *For every formula $\varphi \in \Delta_0$ and x-finite $\vec{y}$,*

$$(\forall x_1)(\exists x_2)\dots(Qx_n)\varphi(x_1,\dots,x_n,\vec{y}) \tag{3.118}$$

is equivalent to

$$(\forall^{x\text{-}st}x_1)(\exists^{x\text{-}st}x_2)\dots(Q^{x\text{-}st}x_n)\varphi(x_1,\dots,x_n,\vec{y}). \tag{3.119}$$

Depending on whether n is even or odd, '(Qx_n)' is '$(\exists x_n)$' or '$(\forall x_n)$'.

The following lemma greatly reduces the number of applications of transfer in a proof. We sometimes refer to it as the 'transfer lemma'.

3.7. Lemma. *For every formula $\varphi \in \Delta_0$ and x-finite $\vec{y}$, if Σ_n-TRANS is available,*

$$(\exists x_1)(\forall x_2)\dots(Qx_n)\varphi(x_1,\dots,x_n,\vec{y}) \tag{3.120}$$

is equivalent to

$$(\exists^{x\text{-}st}x_1)(\forall x_2)\dots(Qx_n)\varphi(x_1,\dots,x_n,\vec{y}), \tag{3.121}$$

and, for $y \gg x$, to

$$(\exists^{y\text{-}st}x_1)(\forall^{y\text{-}st}x_2)\dots(Q^{y\text{-}st}x_n)\varphi(x_1,\dots,x_n,\vec{y}). \tag{3.122}$$

Proof. The equivalence between (3.120) and (3.122) follows from Σ_n-TRANS and the implication '(3.121) $\to$ (3.120)' is trivial. For the implication '(3.120) $\to$ (3.121)', by Σ_n-transfer, (3.120) implies (3.117). Fix x-finite x_1' such the following formula $(\forall^{x\text{-}st}x_2)\dots(Q^{x\text{-}st}x_n)\varphi(x_1',\dots,x_n,\vec{y})$ holds and apply Π_{n-1}-transfer. The resulting formula implies (3.121). □

2. The reduction theorem

In this section, we describe a procedure which reduces a Σ_n-formula with x-standard parameters to a Δ_0-formula. The resulting formula is equivalent to the original one, if Σ_n-TRANS is available. Thus, the following theorem is proved in the theory $I\Delta_0 + \mathrm{NS} + \Sigma_n$-TRANS.

3.8. Theorem. *For $\varphi \in \Delta_0$ and x-standard $\vec{y}$, the formula*

$$(\exists x_1)(\forall x_2)\dots(Qx_n)\varphi(x_1,\dots,x_n,\vec{y}) \tag{3.123}$$

is equivalent to

$$(\exists x_1 \le c_1)(\forall x_2 \le c_2)\dots(Qx_n \le c_n)\varphi(x_1,\dots,x_n,\vec{y}), \tag{3.124}$$

whenever $x \ll c_1 \ll \dots \ll c_n$.

Proof. Let φ, x and $\vec{y}$ be as stated and fix numbers c_i such that $x \ll c_1 \ll \dots \ll c_n$. For better readability, we suppress the x-standard parameters $\vec{y}$ in φ. We first prove the implication '(3.123) $\to$ (3.124)'. Assume n is even. The case for odd n is treated below. From

$$(\exists x_1)(\forall x_2)\dots(\forall x_n)\varphi(x_1,\dots,x_n), \tag{3.125}$$

there follows, by the transfer lemma,

$$(\exists^{x\text{-}st} x_1)(\forall x_2)(\exists x_3)\dots(\forall x_n)\varphi(x_1,x_2,\dots,x_n).$$

As $x \ll c_1$, this implies

$$(\exists x_1 \le c_1)(\forall x_2 \le c_2)(\exists x_3)(\forall x_4)\dots(\forall x_n)\varphi(x_1,x_2,x_3\dots,x_n). \tag{3.126}$$

Fix suitable $x_1' \le c_1$ such that for all $x_2' \le c_2$ there holds

$$(\exists x_3)(\forall x_4)(\exists x_5)\dots(\forall x_n)\varphi(x_1',x_2',x_3\dots,x_n).$$

This formula is in Σ_{n-2}. Repeat the steps that produce (3.126) from (3.125), with $x = c_2$. This yields

$$(\exists x_3 \le c_3)(\forall x_4 \le c_4)(\exists x_5)\dots(\forall x_n)\varphi(x_1',x_2',x_3\dots,x_n),$$

which implies

$$(\exists x_1 \le c_1)(\forall x_2 \le c_2)(\exists x_3 \le c_3)(\forall x_4 \le c_4)(\exists x_5)\dots(\forall x_n)\varphi(x_1,\dots,x_n).$$

Now keep repeating the above process until we obtain (3.124).

If n is odd, we apply the same process as in the even case to obtain

$$(\exists x_1 \le c_1)(\forall x_2 \le c_2)\dots(\forall x_{n-1} \le c_{n-1})(\exists x_n)\varphi(x_1,\dots,x_n).$$

Applying Σ_1-transfer to the innermost existential formula yields

$$(\exists x_1 \le c_1)(\forall x_2 \le c_2)\dots(\forall x_{n-1} \le c_{n-1})(\exists^{c_{n-1}-st} x_n)\varphi(x_1,\dots,x_n),$$

and since $c_n \gg c_{n-1}$, this implies (3.124).

For the reverse implication, we treat the case where n is even; the case where n is odd can be treated analogously. In the former case, we have

$$(\exists x_1 \le c_1)(\forall x_2 \le c_2)\dots(\exists x_{n-1} \le c_{n-1})(\forall x_n \le c_n)\varphi(x_1,\dots,x_n).$$

As $c_n \gg c_{n-1}$, this implies, with $\vec{x} = (x_1,\dots,x_n)$ for brevity,

$$(\exists x_1 \le c_1)\dots(\exists x_{n-3} \le c_{n-3})(\forall x_{n-2} \le c_{n-2})(\exists^{c_{n-1}\text{-}st} x_{n-1})(\forall^{c_{n-1}\text{-}st} x_n)\varphi(\vec{x}),$$

and the transfer lemma, applied to the innermost Σ_2-formula, yields

$$(\exists x_1 \le c_1)\dots(\exists x_{n-3} \le c_{n-3})(\forall x_{n-2} \le c_{n-2})(\exists^{c_{n-2}\text{-}st} x_{n-1})(\forall^{c_{n-2}\text{-}st} x_n)\varphi(\vec{x}).$$

As $c_{n-2} \gg c_{n-3}$, this implies

$$(\exists x_1 \le c_1)\dots(\exists^{c_{n-3}\text{-}st} x_{n-3})(\forall^{c_{n-3}\text{-}st} x_{n-2})(\exists^{c_{n-2}\text{-}st} x_{n-1})(\forall^{c_{n-2}\text{-}st} x_n)\varphi(\vec{x}),$$

Again applying the transfer lemma to the innermost Σ_2-formula yields

$$(\exists x_1 \le c_1)\dots(\exists^{c_{n-3}\text{-}st} x_{n-3})(\forall^{c_{n-3}\text{-}st} x_{n-2})(\exists^{c_{n-3}\text{-}st} x_{n-1})(\forall^{c_{n-3}\text{-}st} x_n)\varphi(\vec{x}).$$

Repeating this process until all n quantifiers are exhausted, we obtain

$$(\exists^{c_1\text{-}st} x_1)(\forall^{c_1\text{-}st} x_2)\dots(\exists^{c_1\text{-}st} x_{n-1})(\forall^{c_1\text{-}st} x_n)\varphi(x_1,\dots,x_n),$$

and Σ_n-transfer with $x = c_1$ yields (3.123). $\square$

Theorem 3.8 states that a Σ_n-statement (with x-finite parameters) about all numbers can be reduced to a Δ_0-statement about a certain initial segment. Thus, this theorem is called the 'Σ_n-reduction theorem' or just 'reduction theorem', if the class of formulas is clear from the context. If we interpret '$y \ll z$' as 'z is very large compared to y', then the reduction theorem tells us that a Σ_n-statement about numbers of size at most x can be reduced to a bounded statement if we have access to n-many higher levels of 'largeness'.

The best-known way to remove quantifiers from a formula is by introducing Herbrand or Skolem functions (see [**8**] or [**21**]). However, the predicate $x \sqsubseteq y$ makes it possible to remove all quantifiers *simultaneously* while keeping the newly introduced objects simple. Indeed, in contrast to Skolemization or Herbrandization, the reduction theorem only introduces new constants c_i.

To conclude this section, we point out an application of the reduction theorem in Reverse Mathematics (see [**46**]). In [**32**], Keisler presents a nonstandard version of each of the 'Big Five' theories of Reverse Mathematics. To this end, he formalizes nonstandard arithmetic in second-order arithmetic (see [**32**, §3 and §4]), using Robinson's external view. After formalizing the stratified framework in second-order arithmetic in the same way (in particular, the natural numbers are exactly the 0-finite numbers), we can obtain ACA^- (the comprehension schema for arithmetical formulas without set parameters) with a minimum of comprehension axioms. Indeed, if TRANS is available, the reduction theorem yields that every arithmetical formula with 0-finite parameters is equivalent to a Δ_0-formula. Thus, comprehension for Δ_0-formulas suffices to obtain ACA^-, if TRANS is available. The latter is not a strong requirement, as, by [**32**, Corollary 7.11], TRANS is not a strong schema in the context of ACA_0. It should be noted, however, that in order to work in second-order arithmetic, we have to adopt Robinson's external view of nonstandard mathematics.

3. Approaching Peano Arithmetic

In this section, we obtain lower bounds for the strength of Σ_n-TRANS. First, we prove that Σ_n-TRANS, when added to $I\Delta_0 + \mathrm{NS}$, makes the resulting theory at least as strong as $I\Sigma_n$. Thus, TRANS takes us all the way up from bounded arithmetic to Peano arithmetic.

In arithmetic, the basic operations $+$ and $\times$ are introduced in Robinson's theory Q. To obtain stronger theories, different flavours of induction can be added, like the following schema (see [**8**, **21**]). The set Φ contains formulas in the language L of arithmetic.

3.9. Axiom schema (Φ-IND). *For every formula $\varphi \in \Phi$, there holds*

$$\big[\varphi(0) \wedge (\forall n)(\varphi(n) \to \varphi(n+1))\big] \to (\forall n)\varphi(n). \tag{3.127}$$

The theory $Q + \Sigma_n$-IND is usually denoted $I\Sigma_n$. The union of all these theories is called Peano arithmetic, or PA for short.

3.10. Theorem. *The theory $I\Delta_0 + \mathrm{NS} + \Sigma_n$-TRANS proves Σ_n-IND.*

Proof. Let φ be a Σ_n-formula in the language of arithmetic and assume the antecedent of Σ_n-IND holds for this formula, i.e. we have

$$\varphi(0, \vec{y}) \wedge (\forall n)(\varphi(n, \vec{y}) \to \varphi(n+1, \vec{y})). \tag{3.128}$$

To increase readability, we suppress the parameters $\vec{y}$ in the rest of the proof. It is an elementary verification that we may do this without loss of generality. Also, it is easily proved that Δ_0-MIN is available in $I\Delta_0$ (see e.g. [**8**]). Thus, we can calculate the least n such that $\phi(n)$, if such there are, for all $\phi \in \Delta_0$.

Now suppose there is an n_0 such that $\neg\varphi(n_0)$. By theorem 3.8, there is a Δ_0-formula $\psi(n)$ such that $\neg\varphi(n)$ is equivalent to $\psi(n)$ for $n \leq n_0$. Let n_2 be the least $n \leq n_0$ such that $\psi(n)$. Thus, there holds $\psi(n_2)$ and also $\neg\psi(n_2 - 1)$ if $n_2 > 0$. But $\psi(n_2)$ is equivalent to $\neg\varphi(n_2)$ and by (3.128), there holds $\varphi(0)$. This implies $n_2 > 0$ and hence we have $\neg\psi(n_2 - 1)$, which is equivalent to $\varphi(n_2 - 1)$. But then there holds $\varphi(n_2 - 1) \wedge \neg\varphi(n_2)$, which contradicts (3.128). Hence, $\varphi(n)$ must hold for all n and we have proved (3.127) for Φ equal to Σ_n. $\square$

By the above, the theory $I\Delta_0 + \text{NS} + \text{TRANS}$ is at least as strong as PA. Karel Hrbacek has suggested that the MacDowell-Specker theorem (see [**34**]) implies that $I\Delta_0 + \text{NS} + \text{TRANS}$ is also conservative over PA. The strength of $I\Delta_0 + \text{NS} + \Sigma_n$-TRANS is conjectured to be $B\Sigma_{n+1}$.

Besides induction, there are other ways of axiomatizing arithmetic. In particular, the so-called 'collection' or 'replacement' axiom schemas yield a series of theories similar to $I\Sigma_n$.

3.11. Axiom schema (Φ-REPL). *For every formula $\varphi \in \Phi$, there holds*

$$(\forall x \leq t)(\exists y)\varphi(x, y) \rightarrow (\exists z)(\forall x \leq t)(\exists y \leq z)\varphi(x, y). \tag{3.129}$$

The theory $I\Delta_0 + \Sigma_n$-REPL is usually denoted $B\Sigma_n$. It is well-known that $I\Sigma_{n+1}$ implies $B\Sigma_{n+1}$ and that the latter implies $I\Sigma_n$ (see e.g. [**8**]). Thus, the theories $B\Sigma_n$ also form a hierarchy of Peano arithmetic. Together with these facts, theorem 3.10 implies that $I\Delta_0 + \text{NS} + \Sigma_{n+1}$-TRANS proves Σ_{n+1}-REPL. The following theorem proves this directly.

3.12. Theorem. *The theory $I\Delta_0 + \text{NS} + \Sigma_{n+1}$-TRANS proves Σ_{n+1}-REPL.*

Proof. Let φ be a Σ_{n+1}-formula and assume the antecedent of Σ_{n+1}-REPL holds for this formula, i.e. we have $(\forall x \leq t)(\exists y)\varphi(x, y)$. Again, we suppress most parameters (but not t) to increase readability. Assume $\varphi(x, y)$ is of the form $(\exists x_1)(\forall x_2)\ldots(Qx_{n+1})\phi(x, y, x_1, \ldots, x_{n+1})$, where $\phi \in \Delta_0$. Fix $c_1, \ldots, c_{n+1}$ such that $x \ll c_1 \ll \cdots \ll c_{n+1}$. By theorem 3.8, for all $x \leq t$, the formula $(\exists y)\varphi(x, y)$ is equivalent to

$$(\exists y \leq c_1)(\exists x_1 \leq c_1)(\forall x_2 \leq c_2)\ldots(Qx_{n+1} \leq c_{n+1})\phi(x, y, x_1, \ldots, x_{n+1}),$$

where $t \ll c_1 \ll \ldots \ll c_{n+1}$. Thus, for all $x \leq t$, there are $y', x_1' \leq c_1$ such that

$$(\forall x_2 \leq c_2)\ldots(Qx_{n+1} \leq c_{n+1})\phi(x, y', x_1', x_2, \ldots, x_{n+1}).$$

By the reduction theorem for $x = c_1$, this formula is equivalent to

$$(\forall x_2)\ldots(Qx_{n+1})\phi(x, y', x_1', x_2, \ldots, x_{n+1}),$$

which yields the consequent of Σ_{n+1}-REPL with $z = c_1$. $\square$

Using the appropriate maximization axioms it is possible to make the bound z a t-standard number. It is well-known that such axioms are available in $I\Delta_0$.

4. Reducing Transfer to the reduction theorem

In the third section, we showed that Σ_n-transfer suffices to obtain the Σ_n-reduction theorem. Interestingly, the former is also equivalent to the latter, by theorem 3.15 below. However, we need the following nonstandard tool, provable in $I\Delta_0 + \mathrm{NS}$. Note that x-infinite parameters are allowed in the formula φ.

3.13. THEOREM (Stratified Overflow and Underflow). *Assume $\varphi \in \Delta_0$.*

(1) *If $\varphi(n)$ holds for all x-finite n, it holds for all n up to some x-infinite $\overline{n}$.* ***(overflow)****.*
(2) *If $\varphi(n)$ holds for all x-infinite n, it holds for all n from some x-finite $\underline{n}$ on.* ***(underflow)****.*

PROOF. For the first item, assume $\varphi(n) \in \Delta_0$ holds for all x-finite n. Then calculate the least n_0 such that $\neg\varphi(n_0)$, which must be x-infinite. Define $\overline{n}$ as $n_0 - 1$. Likewise for the second item. □

3.14. COROLLARY. *Assume $\varphi \in \Delta_0$. If $\varphi(n)$ holds for all x-infinite $n \leq n_0$, with n_0 x-infinite, it holds for all $n \leq n_0$ from some x-finite $\underline{n}$ on.*

PROOF. Define $\psi(n)$ as $\varphi(n) \vee n \geq n_0$ and apply underflow. □

In the following, the previous corollary is also referred to as 'underflow'.

3.15. THEOREM. *In $I\Delta_0$+NS, the Σ_n-reduction theorem is equivalent to the transfer principle Σ_n-TRANS.*

PROOF. By theorem 3.8, the inverse implication is immediate. For the forward implication, we proceed by induction on n. For better readability, we suppress the x-standard parameters $\vec{y}$ in both Σ_n-TRANS and the Σ_n-reduction theorem.

For the case $n = 1$, let φ be as in Σ_1-TRANS and assume $(\exists x_1)\varphi(x_1)$. By the reduction theorem, we have $(\exists x_1 \leq c_1)\varphi(x_1)$, for all $c_1 \gg x$. By underflow, there holds $(\exists^{x\text{-}st} x_1)\varphi(x_1)$. This proves the downward implication in Σ_1-TRANS, i.e. that (3.116) implies (3.117) for $n = 1$. The upward implication is trivial and this case is done.

For the case $n = 2$, let φ be as in Σ_2-TRANS and assume $(\exists x_1)(\forall x_2)\varphi(x_1, x_2)$. By the reduction theorem, we have $(\exists x_1 \leq c_1)(\forall x_2 \leq c_2)\varphi(x_1, x_2)$, for all $c_2 \gg c_1 \gg x$. Fix c_2' and c_1' such that $c_2' \gg c_1' \gg x$. For all x-infinite $d \leq c_1'$, there holds $(\exists x_1 \leq d)(\forall x_2 \leq c_2')\varphi(x_1, x_2)$. By underflow, there is an x-finite $\underline{d}$ such that $(\exists x_1 \leq \underline{d})(\forall x_2 \leq c_2')\varphi(x_1, x_2)$. As $c_2' \gg x$, this implies $(\exists^{x\text{-}st} x_1)(\forall^{x\text{-}st} x_2)\varphi(x_1, x_2)$. This proves the downward implication in Σ_2-TRANS, i.e. that (3.116) implies (3.117) for $n = 2$. The upward implication is easily proved using Σ_1-TRANS, obtained earlier.

For the case $n > 2$, let φ be as in Σ_n-TRANS and assume (3.116) holds. By the Σ_n-reduction theorem, (3.124) follows, for all $c_1, \ldots, c_n$ such that $x \ll c_1 \ll \ldots \ll c_n$. Now fix $c_1', \ldots, c_n'$ such that $x \ll c_1' \ll \ldots \ll c_n'$. For all x-infinite $d \leq c_1'$, there holds

$$(\exists x_1 \leq d)(\forall x_2 \leq c_2') \ldots (Qx_n \leq c_n')\varphi(x_1, \ldots, x_n),$$

and underflow implies $(\exists^{x\text{-}st} x_1)(\forall x_2 \leq c_2') \ldots (Qx_n \leq c_n')\varphi(x_1, \ldots, x_n)$. Fix suitable x-finite x_1' such that for all x-finite x_2', we have

$$(\exists x_3 \leq c_3')(\forall x_4 \leq c_4') \ldots (Qx_n \leq c_n')\varphi(x_1', x_2', x_3, x_4, \ldots, x_n). \tag{3.130}$$

By the Σ_{n-2}-reduction theorem, (3.130) becomes

$$(\exists x_3)(\forall x_4)\ldots(Qx_n)\varphi(x_1', x_2', x_3, x_4, \ldots, x_n). \tag{3.131}$$

By the induction hypothesis, the Σ_{n-2}-reduction theorem yields Σ_{n-2}-TRANS, and Σ_{n-2}-transfer applied to (3.131) yields

$$(\exists^{x\text{-}st} x_3)(\forall^{x\text{-}st} x_4)\ldots(Q^{x\text{-}st} x_n)\varphi(x_1', x_2', x_3, x_4, \ldots, x_n).$$

This can be done for all x-standard x_2' and thus we obtain (3.117). This settles the downward implication in Σ_n-TRANS, i.e. that (3.116) implies (3.117). The upward implication is easily proved using Σ_{n-1}-TRANS, which is available thanks to the induction hypothesis. $\square$

Thus, we know that the reduction theorem is fundamentally connected to PA (and hence ACA_0). In particular, we may 'iterate' a formula in the following way. Let $\varphi(n, X)$ be a Σ_n-formula where X is a variable for a subformula of φ

$$\varphi^0(n) := \varphi(n, 0 = 1) \text{ and } \varphi^{m+1}(n) := \varphi(n, \varphi^m(n)).$$

For each m, the formula $\varphi^m(n)$ is arithmetical. The fourth theory of the Big Five, ATR_0, deals with similar, but *transfinite*, iterations. It would be an interesting challenge to study ATR_0 from the point of view of relative arithmetic.

In conclusion, we poin to [**4**, D.8] where Paris and Harrington formulate the first reasonably natural example of a combinatorial statement that is not provable in Peano Arithmetic. To obtain this famous unprovability result, Paris and Harrington make use of 'indiscernible' numbers which share some properties with the numbers at different levels of infinity in $I\Delta_0 + \mathrm{NS} + \mathrm{TRANS}$. In particular, compare the reduction theorem to [**4**, Claim 2.4].

5. Arithmetical truth

In this section, we investigate the so-called 'truth predicate' or 'truth definition' $\mathbb{T}$ in our stratified framework. This unary predicate has the property that

$$\psi \leftrightarrow \mathbb{T}(\ulcorner\psi\urcorner), \text{ for all sentences } \psi. \tag{T}$$

Thus, the formula $\mathbb{T}(\ulcorner\psi\urcorner)$ simply expresses that ψ is true (or false). As truth is one of the fundamental properties of logic, such predicate $\mathbb{T}$ is a most interesting object of study. For instance, in $I\Sigma_{n+1}$, there is a truth predicate for Σ_n-sentences which respects the logical connectives and this allows for a smooth proof of $I\Sigma_{n+1} \vdash \mathrm{Con}(I\Sigma_n)$ (see [**8**, p. 137]). However, by Tarski's well-known theorem on the undefinability of truth, there is no arithmetical formula $\mathbb{T}$ with the property (T) for *all* arithmetical sentences. Nonetheless, by the reduction theorem, the truth of an arithmetical formula (with x-standard parameters) is equivalent to that of a bounded formula and the truth of the latter can be expressed quite easily. Based on this heuristic idea, we shall obtain an external, i.e. non-arithmetical, formula $\mathbb{T}$ with the property (T) for all arithmetical sentences.

3.16. Theorem. *In $I\Delta_0 + \mathrm{NS} + \mathrm{TRANS}$, there is a truth definition for all arithmetical sentences.*

Proof. By theorem 3.10, $I\Delta_0 + \mathrm{NS} + \Sigma_n$-TRANS is at least as strong as $I\Sigma_n$ and thus the exponential function is available. Hence, we may assume without loss of generality that blocks of existential and universal quantifiers are coded into single quantifiers. In particular, if c is a code for a vector $(c_1, \ldots, c_n)$, then the projection function $[x]_y$ is defined as $[c]_i = c_i$ for $1 \leq i \leq n$. Furthermore, following Buss' arithmetization of metamathematics (see [**8**, Chapter II]), we may assume that the predicate '$\mathrm{Form}_{\Sigma_n \cup \Pi_n}(x)$' which is true if and only if x is the Gödel number of either a Σ_n or Π_n-formula, is available. Now define the predicate $\mathbb{BF}(x, y, c, n)$ as follows. If x is the Gödel number of the $\Sigma_n \cup \Pi_n$-formula

$$(Qx_1)(Qx_2)\ldots(Qx_k)\varphi(x_1, \ldots, x_k, \vec{y}),$$

with $k \leq n$ and y is the Gödel number of a vector $\vec{z}$ with the same length as $\vec{y}$, then $\mathbb{BF}(x, y, c, n)$ is defined as true if

$$(Qx_1 \leq [c]_1)(Qx_2 \leq [c]_2)\ldots(Qx_k \leq [c]_k)\varphi(x_1, \ldots, x_k, \vec{z}).$$

Define $\mathbb{BF}(x, y, c, n)$ as 'false' otherwise. As $I\Delta_0 + \exp$ has a truth definition for Δ_0-formulas, it is clear that the predicate $\mathbb{BF}(x, y, c, n)$ is available. Now define the formula $\mathbb{T}(x, y)$ as

$$(\exists c)(\exists n)\big[\mathrm{Form}_{\Sigma_n \cup \Pi_n}(x) \wedge y = [c]_0 \wedge (\forall i \leq n)([c]_i \ll [c]_{i+1}) \wedge \mathbb{BF}(x, y, c, n)\big]. \quad (3.132)$$

By the reduction theorem, the arithmetical sentence $\psi(\vec{z})$ is true if and only if $\mathbb{T}(\ulcorner \psi \urcorner, \ulcorner \vec{z} \urcorner)$. □

As formula (3.132) explicitly involves the predicate '$\ll$', Tarski's theorem does not contradict the previous corollary. Indeed, the reduction theorem does not apply to external formulas and thus the usual diagonalization argument does not go through.

In Latin, 'infinite' literally means 'the absence of limitation'. In the stratified framework, where the 'infinite' abounds, there is indeed no limitation to our knowledge of arithmetical truth.

6. Philosophical considerations

In the final section, we argue that the reduction theorem yields a formalization of Nelson's notion of impredicativity (see [**37**]). The latter is a key ingredient of Nelson's philosophy of mathematics, which is described by Buss as 'radical constructivism' (see [**9**]).

In Nelson's philosophy, there is no finished set of natural numbers. The only numbers that 'exist' for him, are numbers which have been constructed (thus, finitely many, at any given time). By rejecting the 'platonic' existence of the natural numbers as a finished totality, the induction principle also becomes suspect. This is best expressed in the following quote by Nelson ([**37**, p. 1]).

> The reason for mistrusting the induction principle is that it involves an impredicative concept of number. It is not correct to argue that induction only involves the numbers from 0 to n; the property of n being established may be a formula with bound variables that are thought of as ranging over all numbers. That is, the induction principle assumes that the natural number system is given. A number is conceived to be an object satisfying every inductive formula; for a particular inductive formula, therefore, the bound variables are conceived to range over objects satisfying every inductive formula, including the one in question.

As an example, take Σ_1-induction as in (3.127) where $\varphi(n)$ is $(\exists m)\psi(m,n)$, with $\psi \in \Delta_0$. Even if n only ranges over numbers that have been constructed so far, the existential quantifier $(\exists m)$ may refer to numbers that have not been defined at this point. For this reason, Σ_1-induction is considered meaningless by Nelson. In general, any statement that potentially refers to numbers that have not been defined at that point, is called 'impredicative' and Nelson only deems predicative (i.e. not impredicative) mathematics to be meaningful. Next, we attempt to formalize this notion of impredicativity. As is to be expected, such formalization requires us to step outside of predicative mathematics.

We work in $I\Delta_0 + \mathrm{NS} + \Sigma_1$-TRANS. According to Nelson, there are only finitely many numbers available at any given time. Thus, assume that all numbers that are available at this moment in predicative arithmetic are x-finite, for some x. Now consider the following induction axiom, which is essentially Σ_1-IND for ψ, limited to x-finite numbers,

$$\big[(\exists n)\psi(n,0) \wedge (\forall^{x\text{-}st} m)((\exists n)\psi(n,m) \to (\exists n)\psi(n,m+1))\big] \to (\forall^{x\text{-}st} m)(\exists n)\psi(n,m). \tag{3.133}$$

Here, ψ is in Δ_0 and the possible x-standard parameters have been surpressed. Fix a number $c \gg x$. In $I\Delta_0 + \mathrm{NS} + \Sigma_1$-TRANS, (3.133) is equivalent to

$$\begin{aligned}\big[(\exists n \le c)\psi(n,0) \wedge (\forall^{x\text{-}st} m)((\exists n \le c)\psi(n,m) \to &(\exists n \le c)\psi(n,m+1))\big] \\ &\to (\forall^{x\text{-}st} m)(\exists n \le c)\psi(n,m).\end{aligned}$$

Although induction for bounded formulas is acceptable in predicative arithmetic, the previous formula is not: the bound c used to bound '$(\exists n)$' is not x-finite and hence this number is not available in predicative arithmetic yet. Thus, we see that in $I\Delta_0 + \mathrm{NS} + \Sigma_1$-TRANS, the limited Σ_1-induction axiom (3.133) indeed refers to numbers which are not available at this point in predicative mathematics and as such, Σ_1-IND is not acceptable in the latter. Again, we stress that the previous steps take us outside of predicative arithmetic, i.e. the formalization of impredicativity goes beyond predicative arithmetic.

Obviously, this generalizes to Σ_n-induction, for all $n \in \mathbb{N}$. However, Σ_{n+1}-induction is also impredicative (in the sense of Nelson) 'relative' to Σ_n-induction. Indeed, fix numbers $x \ll c_1 \ll \cdots \ll c_{n+1}$. By the reduction theorem, a Σ_{n+1}-formula (with x-finite parameters) is equivalent to a Δ_0-statement about numbers below c_{n+1}, whereas a Σ_n-formula (with x-finite parameters) is equivalent to a Δ_0-statement about numbers below c_n. Hence, both Σ_n-IND and Σ_{n+1}-IND, limited to x-standard numbers, can be written in a similar equivalent form as the previous centered formula. Thus, even if we regard this limited form of Σ_n-induction (and hence all numbers below c_n) as 'basic', the limited form of Σ_{n+1}-induction refers to numbers which are not basic, namely c_{n+1}.

In light of the above, we may also interpret $x \sqsubseteq y$ as 'x is available when y is'. This interpretation makes the impredicative character of induction apparent.

3.17. Acknowledgement. I wish to thank Karel Hrbacek for his valuable advice, comments and suggestions with regard to this chapter. The contents of this chapter is published in [**43**], in honour of our youngest family member Dario Pleysier (born on January 11, 2010).

APPENDIX A

Technical Appendix

In this appendix, we carry out the 'bootstrapping' process for ERNA, mentioned in section 2.1.3, in full detail.

1. Fundamental functions of ERNA

For further use we collect here some definable functions, being terms of the language that (provably in ERNA) have the properties of the function.

((i)) The identity function $id(x) = x$ is $\pi_{1,1}$.

((ii)) For each constant τ and each arity k, the function

$$C_{k,\tau}(x_1, \dots, x_k) = \tau,$$

is $\pi_{k+1,k+1}(x_1, \dots, x_k, \tau)$.

((iii)) The hypersequence

$$r(n) = \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{if } n \geq 1 \end{cases}$$

is $\mathrm{rec}^k_{\sigma\tau}$ with $k = 1, \sigma = 0, \tau = C_{2,1}$.

((iv)) The function

$$\zeta(x) = \begin{cases} 1 & \text{if } x = 0 \\ x & \text{otherwise} \end{cases}$$

is $1 + x - r(\lceil |x| \rceil)$.

((v)) The functions

$$h(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

are $\frac{x+|x|}{2\zeta(x)}$ and $\frac{1}{2} + \frac{\zeta(|x|)}{2\zeta(x)}$, respectively.

((vi)) For constants $a < b$, the function

$$1_{(a,b]}(x) = \begin{cases} 1 & \text{if } a < x \leq b \\ 0 & \text{otherwise} \end{cases}$$

is $h(x-a)H(b-x)$. Likewise for the characteristic functions of $[a, b]$, (a, b) and $[a, b)$.

((vii)) For constants $a < b$ and terms $\rho(x), \sigma(x), \tau(x)$, the function

$$d_{a,b,\rho,\sigma,\tau}(x) = \begin{cases} \sigma(x) & \text{if } a < x \leq b \text{ and } \rho(x) > 0 \\ \tau(x) & \text{otherwise} \end{cases} \tag{A.134}$$

is $1_{(a,b]}(x)(h(\rho(x))\,\sigma(x) + (1 - h(\rho(x)))\,\tau(x))$. Likewise for $a < x < b$, $a \leq x \leq b$ and $a \leq x < b$ and/or $\rho(x) < 0$, $\rho(x) \leq 0$ and $\rho(x) \geq 0$. Any such construction will be called a *definition by cases*. The interval may be omitted; if so, ρ, σ, τ in $d_{\rho,\sigma,\tau}$ are allowed to have more than one free variable.

((viii)) The function

$$\max_\tau(n) = \begin{cases} \tau(0) & \text{if } n = 0 \\ \begin{cases} \tau(n) & \text{if } \max_\tau(n-1) < \tau(n) \\ \max_\tau(n-1) & \text{if } \max_\tau(n-1) \geq \tau(n) \end{cases} & \text{if } n > 0 \end{cases}$$

introduced in [**49**] computes the greastest of all $\tau(m)$ for $m \leq n$.

((ix)) $\mathrm{least}_\tau(n) := -\max_{-\tau}(n)$ computes the least of all $\tau(m)$ for $m \leq n$.

((x)) The function $\mathrm{even}(n) := H(n/2 - \lceil n/2 \rceil)$ decides whether a hypernatural is even or not; likewise for $\mathrm{odd}(n)$.

Defining summation and product operators requires the following lemma. Its proof relies on hypernatural induction; this explains why $f(n)$ must be internal.

A.1. Lemma. *Let $f(n)$ be an internal hypersequence, defined for all hypernatural numbers, and not involving ω or* min. *If $\|f(n)\| \leq 2^n_k$ ($k \in \mathbb{N}$), and $g(n)$ is the unary term $\mathrm{rec}^{k+2}_{\sigma\tau}$ obtained from the terms $\sigma = f(0)$ and $\tau(n, x) = f(n) + x$, then*

$$g(n) \text{ is defined and } \|g(n)\| \leq 2^n_{k+2}. \tag{A.135}$$

Proof. First, it is easily verified by induction that $2n < 2^n$ for $n \geq 3$. In particular we have for $n \geq 3$ that $n < 2^n$ and $n+3 < 2^n$, hence $n(n+3) < 2^{2n} < 2^{2^n}$. As the inequality $n^2 + 3n \leq 2^{2^n}$ is also valid for $n = 0, 1, 2$, it holds for all n.

Next,

$$2^{2n} \left(2^n_k\right)^{n+1} \leq 2^n_{k+2} \tag{A.136}$$

for all hypernatural numbers $n \geq 0$ and natural $k \geq 1$. For $k = 1$ the statement reduces to

$$2^{n^2+3n} \leq 2^{2^{2^n}},$$

which holds by the last inequality obtained. Now, supposing (A.136) valid for all hypernatural numbers up to k, we estimate $2^{2n}\left(2^n_{k+1}\right)^{n+1}$. The first factor being less than the second, the product is at most $\left(2^n_{k+1}\right)^{2n+2}$, i.e. $2^{(2n+2)\,2^n_k}$. An easy induction shows that $2n + 2 \leq 2^{2n}$ for $n \geq 1$. Using this in the last estimate, we get

$$2^{2n}\left(2^n_{k+1}\right)^{n+1} \leq 2^{2^{2n}\,2^n_k} \leq 2^{2^{2n}\,(2^n_k)^{n+1}},$$

also for $n = 0$. By the induction hypothesis, the upper bound is at most $2^{2^n_{k+2}}$, i.e. 2^n_{k+3}. This concludes the inductive proof of (A.136).

It follows from (A.136) that the statement

$$g(n) \text{ is defined and } \|g(n)\| \leq 2^{2n}\left(2^n_k\right)^{n+1} \tag{A.137}$$

is stronger than (A.135). We now prove it by hypernatural induction. For $n = 0$ it reduces to $g(0) \neq\uparrow$ and $\|g(0)\| \leq 1$. Since $g(0) = f(0)$ by axiom schema 1.31, we are left with $\|f(0)\| \leq 2^0_k$, which is the very assumption for $n = 0$. Next, assume that (A.137) is valid for hypernatural numbers up to n. By axiom schema 1.31 we know that $g(n+1)$ equals $g(n) + f(n+1)$ if this expression is defined and its weight does not exceed 2^{n+1}_{k+2}. As $g(n)$ and $f(n+1)$ are assumed to be defined, their sum also is. Its weight can be estimated from theorem 1.27.4, which implies that

$\|x+y\| \leq 4\|x\|\,\|y\|$ if $\|x\| \geq 1$ and $\|y\| \geq 1$. Both $g(n)$ and $f(n+1)$, being defined, have weight ≥ 1. Hence

$$\|g(n) + f(n+1)\| \leq 2^{2n+2}\,(2_k^n)^{n+1}\,2_k^{n+1} \tag{A.138}$$

by the assumptions on the weights of $g(n)$ and $f(n)$. Increasing $(2_k^n)^{n+1}$ to $\left(2_k^{n+1}\right)^{n+1}$ yields the upper bound $2^{2n+2}\left(2_k^{n+1}\right)^{n+2}$. Therefore (A.138) implies that

$$\|g(n+1)\| \leq 2^{2n+2}\left(2_k^{n+1}\right)^{n+2},$$

which concludes the inductive proof of (A.137). □

A.2. Lemma. *If, in the previous lemma, an estimate $\|f(n)\| \leq 2_{k'}^n$ is used to obtain a term $g' := \mathrm{rec}_{\sigma\tau}^{k'+2}$ instead of $g := \mathrm{rec}_{\sigma\tau}^{k+2}$, then $g'(n) = g(n)$ for all hypernatural numbers n.*

Proof. As we verified in the previous lemma, $g(n+1) = g(n) + f(n+1)$. Likewise, we have $g'(n+1) = g'(n) + f(n+1)$. These equations imply a straightforward hypernatural induction. □

A.3. Notation. For an internal term $f(n)$, defined for all hypernatural numbers, and not involving ω or min, we write

$$\sum_0^n f$$

for the unary term $g(n)$ obtained in lemma A.1. It follows from lemma A.2 that this term is independent from the estimate on $\|f\|$. For convenience, we shall also use dummy variables, e.g. $\sum_{i=0}^n f(i)$.

To add an extra free variable giving the lower limit, put

$$\sum_m^n f = \begin{cases} \sum_0^n f & \text{if } m = 0 \\ \sum_0^n f - \sum_0^{m-1} f & \text{if } 0 < m \leq n \\ \uparrow & \text{otherwise.} \end{cases}$$

Starting from a term $f(n, \vec{x})$ with arity > 1 results in

$$\sum_{l=m}^n f(l, \vec{x}),$$

whose weight is $\leq 2_{k+2}^{\|n,\vec{x}\|}$ if $\|f(n,\vec{x})\| \leq 2_k^{\|n,\vec{x}\|}$ $(k \in \mathbb{N})$.

The estimates which theorem 1.27 gives for $\|xy\|$ are the same as those for $\|x+y\|$. Therefore, all of the preceding can be repeated to yield a product operator $\prod$ alongside $\sum$.

2. Applications of fundamental functions

We now use $\sum$ and $\prod$ to equip ERNA with pairing functions, used to reduce multivariable formulas to single-variable ones. To encode the couple (n, m) into a unique hypernatural k, set

$$k = 2^n(2m+1) - 1.$$

For the inverse operation, set

$$m = \begin{cases} k/2 & \text{if even}(k) \\ \frac{1}{2}\sum_{l=1}^{k}\left(\left(\frac{k+1}{2^l}-1\right)\ \text{odd}\left(\frac{k+1}{2^l}\right)\ \left(1-\prod_{j=0}^{l-1}\text{odd}\left(\frac{k+1}{2^j}\right)\right)\right) & \text{otherwise} \end{cases}$$

and

$$n = \sum_{l=1}^{k}\left(l\ \text{odd}\left(\frac{k+1}{2^l}\right)\ \left(1-\prod_{j=0}^{l-1}\text{odd}\left(\frac{k+1}{2^j}\right)\right)\right).$$

Iterating, we can encode and subsequently decode any finite list of hypernaturals. Thus, we can prove the following multivariable form of transfer, not restricted to hypernatural variables. Before we can use $\sum$ and $\prod$ to resolve bounded quantifiers, we need the following theorem, interesting in its own right.

A.4. THEOREM. *For every internal quantifier-free formula $\varphi(\vec{x})$ not involving* min *or* $\uparrow$*,* ERNA *has a function $T_\varphi(\vec{x})$ such that*

$$\varphi(\vec{x}) \textit{ is true if and only if } T_\varphi(\vec{x}) = 1$$
$$\varphi(\vec{x}) \textit{ is false if and only if } T_\varphi(\vec{x}) = 0.$$

PROOF. Given such a formula $\varphi(\vec{x})$, resolve every occurrence of $\to$, leaving only the logical symbols $\wedge, \vee, \neg$. The proof will be completed using induction on the total number N of occurrences of these symbols. If $N = 0$, the formula is atomic and, being internal, the form $\tau_1(\vec{x}) \approx \tau_2(\vec{x})$ is excluded. Three possible types remain to be considered. In defining the corresponding formula T_φ we use ERNA's function $d_{\rho\sigma\tau}$ defined in (A.134). For $\tau_1(\vec{x}) \le \tau_2(\vec{x})$, take $d_{\tau_2-\tau_1,1,0}(\vec{x})$; for $\tau_1(\vec{x}) = \tau_2(\vec{x})$: $d_{\tau_2-\tau_1,1,0}(\vec{x})\, d_{\tau_1-\tau_2,1,0}(\vec{x})$; finally, for $\mathcal{N}(\tau(\vec{x}))$: $d_{\lceil\tau\rceil-\tau,1,0}(\vec{x})\, d_{\tau-\lceil\tau\rceil,1,0}(\vec{x})\, d_{\tau,1,0}(\vec{x})$, which expresses that $\lceil\tau(\vec{x})\rceil = \tau(\vec{x})$ and $\tau(\vec{x}) \ge 0$.

Next, assume the theorem holds for all formulas $\psi, \phi, \ldots$ with N occurrences of $\vee, \wedge$ and $\neg$, and consider a formula with one occurrence more. For $\neg\psi(\vec{x})$, take $1-T_\psi(\vec{x})$; for $\psi(\vec{x}) \wedge \phi(\vec{x})$: $T_\psi(\vec{x})T_\phi(\vec{x})$, and for $\psi(\vec{x}) \vee \phi(\vec{x})$: $T_\psi(\vec{x}) + T_\phi(\vec{x}) - T_\psi(\vec{x})T_\phi(\vec{x})$. □

Essentially, the same result is also proved for the reduced Chuaqui and Suppes system NQA$^-$ in lemma 2.4 of [**42**]. Both proofs can easily be translated from one theory to the other.

For certain $\vec{x}$, the formula $\varphi(\vec{x})$ may be neither true nor false, for instance $1/x > 0$ for $x = 0$. We will tacitly assume that all formulas used have been adapted to exclude such 'critical points'.

A.5. COROLLARY. *For every pair of terms $\sigma(\vec{x}), \tau(\vec{x})$ and every internal quantifier-free formula $\varphi(\vec{x})$ not involving* min *or* $\uparrow$*,* ERNA *has a function*

$$d_{\varphi\sigma\tau}(\vec{x}) = \begin{cases} \sigma(\vec{x}) & \textit{if } \varphi(\vec{x}) \\ \tau(\vec{x}) & \textit{otherwise.} \end{cases} \tag{A.139}$$

PROOF. Apply definition by cases, as described in (A.134), to $\rho(\vec{x}) := T_\varphi(\vec{x})$. □

From now on, 'definition by cases' will include this extension.

A.6. COROLLARY. *For every internal quantifier-free formula $\varphi(n)$ not involving* min *or* $\uparrow$ *and every hypernatural n_0, the internal formula $(\forall n \le n_0)\varphi(n)$ is equivalent to $\prod_{n=0}^{n_0} T_\varphi(n) > 0$ and, likewise $(\exists n \le n_0)\varphi(n)$ is equivalent to $\sum_{n=0}^{n_0} T_\varphi(n) > 0$.*

Iterating and combining, we see that, as long as its quantifiers apply to bounded hypernatural variables, every internal formula not involving min or $\uparrow$ can be replaced by an equivalent quantifier-free one.

Essentially, the same result is also proved for the reduced Chuaqui and Suppes system NQA^- in lemma 2.4 of [**42**]. Both proofs can easily be translated from one theory to the other, a nice example of the 'kinship' between ERNA and the Chuaqui and Suppes system.

Theorem 1.26 allows us to generalize the preceding corollary as follows.

A.7. Corollary. *For every internal quantifier-free formula $\varphi(x)$ not involving* min *or $\uparrow$ and every hypernatural n_0, the sentences $(\exists x)(\|x\| \leq n_0 \wedge \varphi(x))$ and $(\forall x)(\|x\| \leq n_0 \to \varphi(x))$ are equivalent to quantifier-free ones.*

Next we consider a constructive version of theorem 1.38. Avoiding the use of $\min_\varphi$, it results in functions that can be used in recursion.

A.8. Theorem. *Let $\varphi(n)$ be an internal quantifier-free formula, not involving* min *or $\uparrow$.*

(1) *If $\varphi(n)$ holds for every natural n, it holds for all hypernatural n up to some infinite hypernatural $\overline{n}$* ***(overflow)****.*
(2) *If $\varphi(n)$ holds for every infinite hypernatural n, it holds for all hypernatural n from some natural $\underline{n}$ on* ***(underflow)****.*

Both numbers $\overline{n}$ and $\underline{n}$ are given by explicit ERNA*-formulas not involving* min.

Proof. Suppose $\varphi(n)$ is true for all natural numbers n. The hypernatural

$$\overline{n} := \sum_{n=1}^{\omega} \left(T_\varphi(n) \prod_{k=0}^{n-1} T_\varphi(k) \right) \tag{A.140}$$

is well-defined in ERNA. As $\varphi(n)$ holds for all natural n, $\overline{n}$ is infinite and its very definition shows that $\varphi(n)$ is true for all $n \leq \overline{n}$.

Likewise,

$$\underline{n} := \sum_{n=1}^{\omega} (\omega - n) \left(T_{\neg\varphi}(\omega - n) \left(\prod_{k=0}^{n-1} T_\varphi(\omega - k) \right) \right) \tag{A.141}$$

is well-defined. If there are hypernatural $n \leq \omega$ for which $\neg\varphi(n)$, $\underline{n}$ is the largest of these. Hence, $\underline{n}$ is finite and $\varphi(m)$ holds for all hypernatural $m \geq \underline{n} + 1$. $\square$

This theorem has some immediate consequences.

A.9. Corollary.

Let φ be as in the theorem and assume $n_0 \in \mathbb{N}$.

(1) *If $\varphi(n)$ holds for every natural $n \geq n_0$, it holds for all hypernatural $n \geq n_0$ up to some infinite hypernatural $\overline{n}$, independent of n_0.*
(2) *If $\varphi(n_1, \ldots, n_k)$ holds for all natural $n_1, \ldots, n_k$, it holds for all hypernatural $n_1, \ldots, n_k$ up to some infinite hypernatural $\overline{n}$.*

In both cases he number $\overline{n}$ is given by explicit an ERNA*-formula not involving* min.

Proof. For (1), take n_0 as lower limit in (A.140); for (2), use k summations and k products. $\square$

Analogous formulas hold for underflow. Overflow also allows us to prove that the rationals are dense in the finite hyperrationals, being ERNA's version of the 'fundamental theorem of nonstandard analysis'.

A.10. THEOREM. *For every finite a and every natural n there is a rational b such that* $|a-b| < \frac{1}{n}$.

PROOF. If the stament is false, there exists a finite number a_0 and a natural n_0 such that $|a_0 - b| \geq \frac{1}{n_0}$ for all rational b. Then

$$(\forall b)\left(\|b\| \leq n \to |a_0 - b| \geq \frac{1}{n_0}\right) \tag{A.142}$$

for all natural n. By corollary A.7, this formula is equivalent to a quantifier-free formula, and by theorem A.8, we can apply overflow. Hence, (A.142) continues to hold for n up to some infinite ω_1. Set

$$\omega_2 = \left\lfloor \frac{\omega_1}{\lfloor a_0 \rfloor + 1} \right\rfloor$$

and divide the interval $\big[\lfloor a_0 \rfloor, \lceil a_0 \rceil\big]$ in subintervals of length $\frac{1}{\omega_2} \approx 0$. All points in $\big[\lfloor a_0 \rfloor, \lceil a_0 \rceil\big]$, in particular a_0, are infinitely close to a number of the grid. For $m \leq \omega_2$, $\lfloor a_0 \rfloor + \frac{m}{\omega_2}$ is a point of the grid and

$$\left\| \lfloor a_0 \rfloor + \frac{m}{\omega_2} \right\| = \left\| \frac{\lfloor a_0 \rfloor \omega_2 + m}{\omega_2} \right\| \leq \lfloor a_0 \rfloor \omega_2 + m \leq \lfloor a_0 \rfloor \omega_2 + \omega_2 \leq \omega_1.$$

Hence all points of the grid have weight less than ω_1, contradicting (A.142) for $n = \omega_1$. □

The following theorem is the dual of the previous one.

A.11. THEOREM. *In* ERNA*, there are hyperrationals of arbitrarily large weight between any two numbers.*

PROOF. If a is a hyperrational, $\|-a\| = \|a\|$ and $\|1/a\| = \|a\|$ if $a \neq 0$. Hence, we can restrict ourselves to given hyperrationals $1 \leq a < b$. Write $a = \frac{a_1}{a_2}$ with a_1 and a_2 relatively prime hypernaturals. From $a \geq 1$ we deduce that $\|a\| = \max(a_1, a_2) = a_2$. Choose a hypernatural n so large that $a < a + \frac{1}{n} < b$ and $n > a_2$. As Euclid's proof of the infinitude of the prime numbers can easily be formalized in ERNA, we may assume that n is prime. This implies that $a_2 n$ and $a_1 n + a_2$ are relatively prime. Indeed, n is not a common divisor, as it would divide $a_2 < n$. Therefore, a common divisor $d > 1$ would divide a_2, hence also $(a_1 n + a_2) - a_2$ and finally a_1. Therefore

$$\left\|a + \tfrac{1}{n}\right\| = \left\|\tfrac{a_1 n + a_2}{a_2 n}\right\| = \max(a_1 n + a_2, a_2 n) = a_1 n + a_2 = \|a\| n + a_2, \tag{A.143}$$

growing arbitrarily large with n □

A.12. NOTATION. We write $(\forall \omega)\varphi(\omega, \vec{x})$ for $(\forall n)(n \text{ is infinite} \to \varphi(n, \vec{x}))$. Likewise, $(\exists \omega)\varphi(\omega, \vec{x})$ means $(\exists n)(n \text{ is infinite} \wedge \varphi(n, \vec{x}))$.

In the following theorem we establish some useful variants of minimization, which will be used in proving theorem 1.100. Again, they are constructive in avoiding the use of min.

A.13. Theorem.
Let M be a hypernatural and ω_1 an infinite hypernatural. Consider a quantifier-free internal formula $\varphi(n,\vec{x})$ and internal hypersequences $f(n)$ and $g(n)$, none involving min *or* $\uparrow$.

(1) *If there are* natural *n's such that $\varphi(n,\vec{x})$, then* ERNA *has a function $m_\varphi(\vec{x})$, with $\|m_\varphi(\vec{x})\| \leq \omega$, which is the least of these.*
(2) *If there are* hypernaturals *$n \leq M$ such that $\varphi(n,\vec{x})$, then* ERNA *has a function $m_{\varphi,M}(\vec{x})$, with $\|m_{\varphi,M}(\vec{x})\| \leq M$, which is the least of these.*
(3) *If there are infinite hypernaturals $n \leq \omega_1$ such that $\varphi(n,\vec{x})$, then* ERNA *has a function $m_{\varphi,\omega_1}(\vec{x})$ with $\|m_{\varphi,\omega_1}(\vec{x})\| \leq \omega_1$, which is the largest of these.*

The functions m_φ, $m_{\varphi,M}$ and m_{φ,ω_1} are given by explicit ERNA*-functions, not involving* min.

Proof. Set

$$m_\varphi(\vec{x}) = \sum_{n=1}^{\omega} \left(n\ T_\varphi(n,\vec{x}) \ \prod_{k=0}^{n-1} T_{\neg\varphi}(k,\vec{x}) \right),$$

yielding a hypernatural which is at most ω. Likewise for $m_{\varphi,M}$. Finally we use ‘definition by cases’ to obtain m_{φ,ω_1}, which is equal to

$$\sum_{n=1}^{\omega_1} \left((\omega_1 - n)\ T_\varphi(\omega_1 - n,\vec{x}) \ \prod_{k=0}^{n-1} T_{\neg\varphi}(\omega_1 - k,\vec{x}) \right)$$

if $\neg\varphi(\omega_1,\vec{x})$, and equal to ω_1 otherwise. □

The following theorem generalizes overflow to special external formulas.

A.14. Theorem. *Let φ, f and ω_1 be as in the previous theorem.*

(1) *If $f(n)$ is infinite for every $n \in \mathbb{N}$, it continues to be so for all hypernatural n up to some hypernatural number ω_2.*
(2) *If $(\forall^{st} n)(\exists \omega \leq \omega_1)\varphi(n,\omega)$, then there is an infinite hypernatural ω_3 such that $(\forall^{st} n)(\exists \omega \geq \omega_3)\varphi(n,\omega)$.*

Proof. For (1), apply overflow to the formula $f(n) > n$. For (2), let $m_{\varphi,\omega_1}(n)$ be the function obtained by applying theorem A.13.(3) to $(\exists \omega \leq \omega_1)\varphi(n,\omega)$. Then $m_{\varphi,\omega_1}(n)$ is infinite for all $n \in \mathbb{N}$ and by (1), $m_{\varphi,\omega_1}(n)$ is infinite for all $n \leq \omega_2$ for some infinite ω_2. Use 1.((ix)) to obtain the least of these. □

Note that item (2) of the theorem contains $(\forall)(\exists)$, which makes it a Π_2-formula.

APPENDIX B

Dutch Summary

1. Samenvatting

Reverse Mathematics (RM) is een programma in de grondslagen van de wiskunde gesticht door Harvey Friedman in de jaren zeventig ([**17**, **18**]). Het doel van RM is het bepalen van de minimale axioma's $\mathcal{A}$ die een zekere stelling $\mathcal{T}$ uit de 'alledaagse' wiskunde bewijzen. In vele gevallen observeert men dat deze minimale axioma's $\mathcal{A}$ ook equivalent zijn met de stelling $\mathcal{T}$. Een mooi voorbeeld hiervan is gegeven in stelling 1.2. In de praktijk zijn de meeste stellingen uit de alledaagse wiskunde equivalent met één van de vier systemen WKL_0, ACA_0, ATR_0 en Π^1_1-CA_0, ofwel bewijsbaar in de basistheorie RCA_0. Een excellente inleiding tot RM vindt men in Stephen Simpson's boek [**46**]. Nietstandaard analyse speelt een belangrijke rol binnen RM ([**32**, **52**, **53**]).

Een van de open problemen in de literatuur is de RM van de theorie $I\Delta_0+\exp$ ([**46**, p. 406]). In hoofstuk I formuleren we een oplossing voor dit probleem in stelling 1.3. Volgens deze stelling blijven de equivalenties uit stelling 1.2 bewaard indien we gelijkheid vervangen door de infinitesimale gelijkheid '$\approx$' uit de nietstandaard analyse. De basistheorie is nu ERNA, een nietstandaard uitbreiding van $I\Delta_0 +\exp$. Het principle dat met 'Weak König's Lemma' overeenstemt is het universele-overdrachtsprincipe Π_1-TRANS (zie schema 1.57). In het bijzonder kan men stellen dat de RM van ERNA$+\Pi_1$-TRANS een 'kopie op infinitesimalen na' van de RM van WKL_0 is. Dit impliceert dat RM 'robuust' is in de zin dat deze term in statistiek en informatica gebruikt wordt ([**25**, **35**]). Verder bewijzen we toepassingen van onze resultaten voor de theoretische fysica in de vorm van het 'Isomorphism Theorem' (zie stelling 1.106). Deze filosofische zijstap is de eerste toepassing van RM buiten de wiskunde en heeft tot gevolg dat 'de wijze waarop men aan wiskunde doet in de fysica met zich meebrengt dat experimenten niet kunnen beslissen of de fysische realiteit continu dan wel discreet is' (zie paragraaf 3.2.4). In de rest van hoofdstuk I beschouwen we de RM van ACA_0 en aanverwante (soms constructieve) systemen.

In hoofdstuk II behandelen we een vormelijk probleem in verband met hoofdstuk I. De RM van ERNA $+\,\Pi_1$-TRANS kan namelijk enkel geformuleerd worden door gebruik te maken van een groot aantal kwantorwisselingen, bijvoorbeeld in de eerste fundamentele stelling van de analyse (zie stelling 1.94). Echter, het nietstandaard raamwerk reduceert normaal gezien het aantal van deze wisselingen. In hoofdstuk II beschouwen we een nieuw nietstandaard raamwerk, 'relatieve nietstandaard analyse' genaamd, dat gebruik maakt van verschillende niveaus van oneindigheid. Deze nieuwe vorm van nietstandaard analyse werd gepionierd door Karel Hrbacek en Yves Péraire ([**22**, **40**]). In 2 breiden we ERNA uit tot een relatieve nietstandaard theorie, $\mathrm{ERNA}^{\mathbb{A}}$. In de nieuwe theorie wordt enige basisanalyse ontwikkeld op een universele manier (4). Daarnaast bestuderen we verschillende overdrachtspincipes en tonen we aan dat het nieuwe 'relatieve' systeem een verfijning is van het oude. We bewijzen eveneens een groot aantal stellingen in ERNA en uitbreidingen die een 'vertaling' zijn van stellingen uit $\mathrm{ERNA}^{\mathbb{A}}$ (5). We bekomen ook verschillende resultaten die verband houden met Andreas Weiermann's fasenovergangenprogramma (5.3). De rode draad doorheen dit hoofdstuk is eveneens geïnspireerd door RM (zie bvb. 5.1).

In hoofdstuk III formuleren we een interne axiomatiek voor de relatieve rekenkunde. De theorieën ERNA en $\mathrm{ERNA}^{\mathbb{A}}$ uit de eerste twee hoofdstukken maken namelijk deel uit van de externe axiomatiek en de interne aanpak is in het algemeen veel eleganter. De eenvoudigere interne formalisering laat ons toe de opmerkelijke reductiestelling (zie stelling 3.8) te bewijzen. Dankzij deze kunnen we een waarheidspredicaat voor arithmetische formules binnen de Peano rekenkunde introduceren (zie stelling 3.16) en Nelson's concept 'impredicativiteit' formaliseren (6).

APPENDIX C

Acknowledgments

1. Thanks

In [**13**, Meditation XVII], it is written that 'No man is an island'. This certainly is true for this dissertation. Many people deserve thanks for contributing to it and I thank them in chronological order.

First of all, there is Chris Impens under whose wings I studied Nonstandard Analysis and wrote my Master's and PhD dissertation on this subject. He also taught me the basics of proper teaching and writing and has even attempted to civilize my Dutch language. For all this, I am greatly indebted to him.

In 2004, I met Karel Hrbacek during a conference in Aveiro, Portugal. His ideas concerning relative infinitesimals greatly appealed to me. Over the years, we have exchanged views and ideas. These interactions have been of great value to this thesis, as is clear from Chapters I, II and III. I owe him a great lot for all this.

In 2006, Andreas Weiermann joined our department. Without his suggestions regarding Reverse Mathematics, Chapter I would not be. The international contacts he made possible are also of immense value. Over the years, a research interaction has developed and Andreas became my second supervisor. I cannot thank him enough for all this.

In 2007, I met Kristina Liefke during the LC2007 in Wrocław, Poland. Apart from being my partner in life, she taught me how to look at mathematics with the eyes of an analytic philosopher. Her loving support drives me and my research forward. Words cannot express my gratitude.

I would also like to thank the various funding entities and organizations which have supported my doctoral research. In no particular order, I thank the FWO Vlaanderen, the ASL, Ghent University, Fonds Professor Wuytack, Tohoku University and the Air Force Office of Scientific Research of the US army.

Finally, I would like to thank my parents, friends and family. Special thanks go to my uncle Hugo Gotink and friend Koen Balemans for the lively discussions concerning topics which transcend the profane world.

Sam Sanders February 16 2010

Bibliography

[1] Jeremy Avigad, *Weak theories of nonstandard arithmetic and analysis*, Reverse mathematics 2001, Lect. Notes Log., vol. 21, ASL, La Jolla, CA, 2005, pp. 19–46.

[2] Jeremy Avigad and Jeremy Helzner, *Transfer principles in nonstandard intuitionistic arithmetic*, Archive for Mathmatical Logic **41** (2002), 581–602.

[3] Jeremy Avigad, *Number theory and elementary arithmetic*, Philos. Math. (3) **11** (2003), no. 3, 257–284.

[4] *Handbook of mathematical logic*, North-Holland Publishing Co., Amsterdam, 1977. Edited by Jon Barwise; Studies in Logic and the Foundations of Mathematics, Vol. 90.

[5] Errett Bishop and Douglas S. Bridges, *Constructive analysis*, Grundlehren der Mathematischen Wissenschaften, vol. 279, Springer-Verlag, Berlin, 1985.

[6] Douglas S. Bridges, *Constructive functional analysis*, Research Notes in Mathematics, vol. 28, Pitman publishing, London, San Francisco and Melbourne, 1979.

[7] Douglas S. Bridges and Luminiţa Simona Vîţă, *Techniques of constructive analysis*, Universitext, Springer, New York, 2006.

[8] Samuel R. Buss, *An introduction to proof theory*, Handbook of proof theory, Stud. Logic Found. Math., vol. 137, North-Holland, Amsterdam, 1998, pp. 1–78.

[9] ———, *Nelson's work on logic and foundations and other reflections on the foundations of mathematics*, Diffusion, quantum theory, and radically elementary mathematics, Math. Notes, vol. 47, Princeton Univ. Press, Princeton, NJ, 2006, pp. 183–208.

[10] Chen Chung Chang and H. Jerome Keisler, *Model theory*, 3rd ed., Studies in Logic and the Foundations of Mathematics, vol. 73, North-Holland, Amsterdam, 1990.

[11] Rolando Chuaqui and Patrick Suppes, *Free-variable Axiomatic Foundations of Infinitesimal Analysis: A Fragment with Finitary Consistency Proof*, Journal of Symbolic Logic **60** (1995), 122-159.

[12] Kevin Davey, *Is mathematical rigor necessary in physics?*, British J. Philos. Sci. **54** (2003), no. 3, 439–463.

[13] John Donne, *From Devotions upon Emergent Occasions*, Available online at `http://www.online-literature.com/donne/409/`, 1624.

[14] Kenneth R. Davidson and Allan P. Donsig, *Real Analysis with Real Applications*, Prentice Hall, Upper Saddle River, N.J., 2002.

[15] Solomon Feferman, *Arithmetization of metamathematics in a general setting*, Fund. Math. **49** (1960/1961), 35–92.

[16] Richard Feynmann, Ralph Leighton (contributor), and Edward Hutchings (editor), *Surely You're Joking, Mr. Feynman!: Adventures of a Curious Character*, W. W. Norton, 1985.

[17] Harvey Friedman, *Some systems of second order arithmetic and their use*, Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 1, Canad. Math. Congress, Montreal, Que., 1975, pp. 235–242.

[18] ———, *Systems of second order arithmetic with restricted induction, I & II (Abstracts)*, Journal of Symbolic Logic **41** (1976), 557–559.

[19] ———, *Grand Conjectures*, FOM mailing list (16 April 1999).

[20] Evgeniĭ I. Gordon, *Nonstandard methods in commutative harmonic analysis*, Vol. 164, American Mathematical Society, Providence, RI, 1997.

[21] Petr Hájek and Pavel Pudlák, *Metamathematics of First-Order Aritmetic*, Springer, 1998.

[22] Karel Hrbacek, *Stratified analysis?*, The strength of nonstandard analysis, Springer Wien, New York, Vienna, 2007, pp. 47–63.

[23] ———, *Relative Set Theory: Internal View*, Journal of Logic and Analysis **1** (2009), no. 8, 1–108.

[24] Karel Hrbacek, Olivier Lessmann, and Richard O'Donovan, *Analysis with ultrasmall numbers*, To appear in American Mathematical Monthly (2010).

[25] Peter J. Huber and Elvezio M. Ronchetti, *Robust statistics*, 2nd ed., Wiley Series in Probability and Statistics, John Wiley & Sons Inc., Hoboken, NJ, 2009.

[26] Albert E. Hurd and Peter A. Loeb, *An introduction to nonstandard real analysis*, Pure and Applied Mathematics, vol. 118, Academic Press Inc., Orlando, FL, 1985.

[27] Chris Impens and Sam Sanders, ERNA *at work*, The strength of nonstandard analysis, Springer Wien, New York, Vienna, 2007, pp. 64–75.

[28] ______, *Transfer and a supremum principle for* ERNA, Journal of Symbolic Logic **73** (2008), 689-710.

[29] ______, *Saturation and* Σ_2*-transfer for* ERNA, Journal of Symbolic Logic **74** (2009), 901-913.

[30] Hajime Ishihara, *Reverse mathematics in Bishop's constructive mathematics*, Philosophia Scientiae (Cahier Spécial) **6** (2006), 43-59.

[31] ______, *Constructive reverse mathematics: compactness properties*, From sets and types to topology and analysis, Vol. 48, Oxford Univ. Press, Oxford, 2005, pp. 245–267.

[32] H. Jerome Keisler, *Nonstandard arithmetic and reverse mathematics*, Bull. Symbolic Logic **12** (2006), no. 1, 100–125.

[33] Ulrich Kohlenbach, *Things that can and things that can't be done in PRA*, Annals of Pure and Applied Logic **102** (2000), 223–245.

[34] Roman Kossak and James H. Schmerl, *The structure of models of Peano arithmetic*, Oxford Logic Guides, vol. 50, The Clarendon Press Oxford University Press, Oxford, 2006. Oxford Science Publications.

[35] The Linux Information Project, *Robustness*, 2005. `http://www.linfo.org/robust.html`.

[36] Edward Nelson, *Internal set theory: a new approach to nonstandard analysis*, Bull. Amer. Math. Soc. **83** (1977), no. 6, 1165–1198.

[37] ______, *Predicative arithmetic*, Mathematical Notes, vol. 32, Princeton University Press, Princeton, NJ, 1986.

[38] ______, *The virtue of simplicity*, The strength of nonstandard analysis, Springer Wien, New York, Vienna, 2007, pp. 27–32.

[39] Vladimir Kanovei and Michael Reeken, *Nonstandard analysis, axiomatically*, Springer, 2004.

[40] Yves Péraire, *Théorie relative des ensembles internes*, Osaka J. Math. **29** (1992), no. 2, 267–297 (French).

[41] Abraham Robinson, *Non-standard analysis*, North-Holland, Amsterdam, 1966.

[42] Michal Rössler and Emil Jeřábek, *Fragment of nonstandard analysis with a finitary consistency proof*, Bulletin of Symbolic Logic **13** (2007), 54-70.

[43] Sam Sanders, *Relative Arithmetic*, Accepted for publication by Mathematical Logic Quarterly (2010).

[44] ______, *More infinity for a better finitism*, Accepted for publication by Annals of Pure and Applied Logic (2010).

[45] Stephen G. Simpson, *Which set existence axioms are needed to prove the Cauchy/Peano theorem for ordinary differential equations?*, Journal of Symbolic Logic **49** (1984), no. 3, 783–802.

[46] ______, *Subsystems of second order arithmetic*, 2nd ed., Perspectives in Logic, Cambridge University Press, Cambridge, 2009.

[47] ______ (ed.), *Reverse mathematics 2001*, Lecture Notes in Logic, vol. 21, Association for Symbolic Logic, La Jolla, CA, 2005.

[48] Keith D. Stroyan and Willem A.J. Luxemburg, *Introduction to the theory of infinitesimals*, Academic Press, 1976.

[49] Richard Sommer and Patrick Suppes, *Finite Models of Elementary Recursive Nonstandard Analysis*, Notas de la Sociedad Mathematica de Chile **15** (1996), 73-95.

[50] ______, *Dispensing with the Continuum*, Journal of Math. Psychology **41** (1997), 3-10.

[51] Willem W. Tait, *Finitism*, The Journal of Philosophy **78** (1981), 524-564.

[52] Kazuyuki Tanaka, *The self-embedding theorem of* WKL_0 *and a non-standard method*, Annals of Pure and Applied Logic **84** (1997), 41-49.

[53] ______, *Non-standard analysis in* WKL_0, Math. Logic Quart. **43** (1997), no. 3, 396–400.

[54] Terence Tao, *Ultrafilters, nonstandard analysis, and epsilon management*, 2007. Wordpress blog, `http://terrytao.wordpress.com/`.

[55] Andreas Weiermann, *Phase transitions for Gödel incompleteness*, Ann. Pure Appl. Logic **157** (2009), no. 2-3, 281–296, DOI 10.1016/j.apal.2008.09.012.

[56] Keita Yokoyama, *Complex analysis in subsystems of second order arithmetic*, Arch. Math. Logic **46** (2007), no. 1, 15–35.

[57] ———, *Standard and non-standard analysis in second order arithmetic*, PhD thesis, Tohoku University, Sendai, 2007. Available online at `http://www.math.tohoku.ac.jp/tmj/PDFofTMP/tmp34.pdf`.

[58] Nobuyuki Sakamoto and Keita Yokoyama, *The Jordan curve theorem and the Schönflies theorem in weak second-order arithmetic*, Arch. Math. Logic **46** (2007), no. 5-6, 465–480.

Created using LaTeX

All results in this dissertation were discovered, not invented.

Thesis version 1.337